

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 ; Search time 4703.97 Seconds
(without alignments)
10155.006 Million cell updates/sec

Title: US-09-596-141C-3

Perfect score: 1643
Sequence: 1 gaaatctctgtgtgctc.....caggcgagcgaggagctc 1643

Scoring table: IDENTITY_NUC
Gapop 10-C, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4-09280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_hugo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1643	100.0	1643	6	AX060715	Sequence
2	1643	100.0	1643	6	AX060894	Sequence
3	1544.8	94.0	96717	9	AL359182	Human DNA
4	1498.6	91.2	183999	6	AX092589	Sequence
5	1485.4	90.4	3231	6	AX351029	Sequence
6	1418	86.3	149034	9	AF275948	Homo sapi
7	1394.4	84.9	2893	6	AX351031	Sequence
8	1247.6	75.9	201144	9	AF287262	Homo sapi
9	1007	61.3	175064	2	AC012230	Homo sapi
10	1006	61.2	1167	9	HA252201	Homo sapi
11	1002.8	61.0	1167	9	AF258623S1	Homo sapi
12	519.8	31.6	69570	2	AC021246	Homo sapi
13	271.6	15.5	69570	2	AC021246	Homo sapi
14	223.8	14.0	90698	2	AC021345	Homo sapi
15	173	10.5	697	9	AF258627	Homo sapi
16	155.4	9.5	186889	2	AL807243	Mus muscu
17	155.4	9.5	278572	10	AF287263	AF287263 Mus muscu
18	153.2	9.3	145833	2	AC125837	AF287263 Rattus no
19	102.2	6.2	2893	6	AX351031	Sequence
20	102.2	6.2	3231	6	AX351029	Sequence
21	99.4	6.0	56717	9	AL359182	Human DNA
22	98	6.0	7260	6	AX253452	Sequence
23	97.8	6.0	49034	9	AF275948	Homo sapi
24	97	5.9	1643	6	AX060715	Sequence
25	97	5.9	1643	6	AX060894	Sequence
26	92.6	5.6	183999	6	AX092589	Sequence
27	91	5.5	221	6	AX351032	Sequence
28	89.4	5.4	1556	9	AK024328	Homo sapi
29	89	5.4	1750	9	AK022254	Homo sapi
30	75	4.6	9854	6	AX127831	Sequence
31	75	4.6	9854	6	AX139818	Sequence
32	69.2	4.2	201144	9	AF287262	Homo sapi
33	67	4.1	10442	6	AX060713	Sequence
34	67	4.1	10442	6	AX060892	Sequence
35	67	4.1	10442	9	AF285167	Sequence
36	60	3.7	10474	6	AX060719	Sequence
37	60	3.7	10474	6	AX060721	Sequence
38	60	3.7	10474	6	AX060898	Sequence
39	60	3.7	10474	6	AX060900	Sequence
40	55.6	3.4	191720	2	AC109982	Rattus no
41	52.6	3.2	125020	9	AF429315	Homo sapi
42	52.2	3.2	125020	9	AF429315	Homo sapi
43	49.6	3.0	96136	2	AC097674	AC097674 Rattus no
44	49.4	3.0	173127	2	AC131142	Rattus no
45	49	3.0	65780	2	AC111741	Rattus no

ALIGNMENTS

RESULT 1
LOCUS AX060715
DEFINITION Sequence 3 from Patent WO078972.
ACCESSION AX060715
VERSION AX060715.1 GI:12406104
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1643)
AUTHORS Lawn, R.M., Wade, D. and Garvin, M.
TITLE Regulation with binding cassette transporter protein abel
JOURNAL Patent: WO 0078972-A 3 28-DEC-2000;

Pred. No. is the number of results predicted by chance to have a

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HG9/Chr9>
RP11-2.7B7 is from the library RP11-11.1 constructed by the group of Pletier de Jong, for further details see

<http://www.ckori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-217B7 it may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-217B7 is at 96717 in this sequence. The true left end of clone RP11-122F10 is at 72980 in this sequence. The true right end of clone RP11-31J20 is at 2000 in this sequence.

FEATURES

```

source
    Location/Qualifiers
        1..96717
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="9"
        /clone="RP11-217B7"
        /clone_lib="RP11-11.1"
        3238..3278
        /note="Sequence from AF275948 sequenced by National
        Institutes of Health, National Heart, Lung and Blood
        Institute, Bethesda, MD 20892, USA."
        34249..84273
        /note="Sequence from overlapping clone RP11-122F10
        (AC026643). Assembly confirmed by restriction digest."
        92050..92163
        /note="Sequence confirmed by restriction digest."
        /note="Sequence from reads from a short insert library
        derived from a single pUC clone. Restriction digest data
        confirm the assembly."
        92411..92557
        /note="Sequence from reads from a short insert library
        derived from a single pUC clone. Restriction digest data
        confirm the assembly."
        misc_feature
            27673 a 21138 c 20380 g 27526 t
    BASE COUNT
    ORIGIN

```

```

Query Match      94.0%; Score 1544.8; DB 9; Length 96717;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 2; Indels 8; Gaps 7;

QY 16  GCTCCACATGACCTCCAGGCGCTCTGG--CCTCTATGGGTCGNCCTGACTGTC 73
Db 18798 GCTCCACATGACCTCCAGGCGCTCTGGGCGCTCTCTATGGGTCGNCCTGAGTGT 18739

QY 74  GATGAACCACTGATGIGAGTACTGGGCTTGAG-CGTGGCCTGGAGATCCTGTTGACTG 132
Db 18738 GATGAACCACTGATGIGAGTACTGGGCTTGAGCCGTGGCCTGGAGATCCTGTTGACTG 18679

QY 133 TAGCTGGAGGGGGCTTGT-CAGCTGAATGCTGTATGAGGCTGGGGAGTTCTGGAAT 191
Db 18678 TAGCATGGAGGGGGCTTGTGTACAGCTGAATGCTGTATGAGGCTGGGGAGTTCTGGAAT 18619

QY 192 ATGATGGAGCTGGAGGTGGGAAGAGTAGGCTTGGGCGAGCTCTCTAIGCCACTCA 251
Db 18618 ATGATGGAGCTGGAGGTGGGAAGAGTAGGCTTGGGCGAGCTCTCTAIGCCACTCA 18559

QY 252 TTCTGCCAAACACTCAGGTCAAACCTGTGAAGAGTCTAAATGTGAATCTGCCCTCAAGGT 311
Db 18558 TTCTGCCAAACACTCAGGTCAAACCTGTGAAGAGTCTAAATGTGAATCTGCCCTCAAGGT 18499

```

```

QY 312  GGCTACAAAGGTATCTTCTCAAGGTAGGAGACC-TGTGGCCTCCACGTGCACTTCCAGG 371
Db 18498 GGCTACAAAGGTATCTTGTCAAGGTAGGAGACCTTGTGGCCTCCACGTGCACTTCCAGG 18439

QY 372  GGCTGCTT-GGGCTCTCTACCGGCTGCTGCCGAGTCTTCTATGAATCT-CCCTTCAGGG 429
Db 18438 GGCTGCTTGGGCTCTCTTACGGGCTGCTGCCGAGTCTTCTATGAATCTGCCCTTCAGGG 18379

QY 430  CAGATTCATATTCAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATAAGGTGACAT 489
Db 18378 CAGATTCATATTCAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATAAGGTGACAT 18319

QY 490  TTAGTTTGTGGCTTGATGAATGACTTAAATATTTAGACATATGGTGCTAGGCTGCAAT 549
Db 18318 TTAGTTTGTGGCTTGATGAATGACTTAAATATTTAGACATATGGTGCTAGGCTGCAAT 18259

QY 550  TGCCTACTCTTGCTTTTGGCTTCCAGTCTGCTGGGAGTTTGGGAGTTTCTGCCCTTACA 609
Db 18258 TGCCTACTCTTGCTTTTGGCTTCCAGTCTGCTGGGAGTTTGGGAGTTTCTGCCCTTACA 18200

QY 610  GCGAAGGCAACAGATAAGTTGGAGTCTGGAGTGGCTACATATTTTACACGACTGCA 669
Db 18199 GCGAAGGCAACAGATAAGTTGGAGTCTGGAGTGGCTACATATTTTACACGACTGCA 18140

QY 670  ATTCTGTGCTGCACTTCACAAATGTATACAACTAAATACAAG-CCTCTGTTTATCA 729
Db 18139 ATTCTGTGCTGCACTTCACAAATGTATACAACTAAATACAAGTCTCTGTTTATCA 18080

QY 730  CAGGAGGCTGATCAATATAAATAAAGGAGGCTGGTCCCAATGTTCTCTGTT 789
Db 18079 CAGGAGGCTGATCAATATAAATAAAGGAGGCTGGTCCCAATGTTCTCTGTT 18021

QY 790  TTGTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 849
Db 18020 TTGTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 17961

QY 850  AGCAGTAAGATGTTCTCTCGGCTCCTCTGAGGAGCTGGGAGCTCAGCTGGGGAATCT 909
Db 17960 AGCAGTAAGATGTTCTCTCGGCTCCTCTGAGGAGCTGGGAGCTCAGCTGGGGAATCT 17901

QY 910  CAAGGCACTAGTTCGCCCTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 969
Db 17900 CAAGGCACTAGTTCGCCCTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 17841

QY 970  CAGCCCATACCCAGAGGACTGTCCGCTTCCCTCACCCTCAGCCCTAGCCCTTGAAGG 1029
Db 17840 CAGCCCATACCCAGAGGACTGTCCGCTTCCCTCACCCTCAGCCCTTGAAGG 17781

QY 1030 AACAAAAAGACAGACAAATGATTTGGCTGCTGAGGAGATTCAGCTAGACTCTCTC 1089
Db 17780 AACAAAAAGACAGACAAATGATTTGGCTGCTGAGGAGATTCAGCTAGACTCTCTC 17721

QY 1090 TCCCCCAATCCCTCCGCTGAGGAACTAAACAAAGGAAAGAAATTCGGGGAAGC 1149
Db 17720 TCCCCCAATCCCTCCGCTGAGGAACTAAACAAAGGAAAGAAATTCGGGGAAGC 17661

QY 1150 AGGATTTAGAGGAAGCAAAATTCACATGGTGGCTGGCTGGCGGAACTGACTAGAG 1209
Db 17660 AGGATTTAGAGGAAGCAAAATTCACATGGTGGCTGGCTGGCGGAACTGACTAGAG 17601

QY 1210 GTCTGGGGGAGCCCGGAGCCCGAGCGCTTCCGCGGCTCTTAGCGCGGCGGCGCGGC 1269
Db 17600 GTCTGGGGGAGCCCGGAGCCCGAGCGCTTCCGCGGCTCTTAGCGCGGCGGCGCGGC 17541

QY 1270 GGGGGAAGGGGAGCGAGAGCCCGGAGCCCTAAGACACTGCTGTACCTCCACCCCAACC 1329
Db 17540 GGGGGAAGGGGAGCGAGAGCCCGGAGCCCTAAGACACTGCTGTACCTCCACCCCAACC 17481

QY 1330 CACCCACCTTCCCCCAACTCCCTAGATGTGTGTGGGCGGCTGAACGTGCGCGCTTAAG 1389
Db 17480 CACCCACCTTCCCCCAACTCCCTAGATGTGTGTGGGCGGCTGAACGTGCGCGCTTAAG 17421

```


QY	1390	GGGCGGGCCCGGCTCCACGTCCTTCTGCTGAGTACGAACACATAAACAAGAGCGG	1449
Db	17420	GGGCGGGCCCGGCTCCACGTCCTTCTGCTGAGTACGAACACATAAACAAGAGCGG	17361
QY	1450	GGAACGGGCGGGGAGGAGACACAGGCTTTGACCGATAGTAACTCTGCGCTCGG	1509
Db	17360	GGAACGGGCGGGGAGGAGACACAGGCTTTGACCGATAGTAACTCTGCGCTCGG	17301
QY	1510	TGCAGCGCAATCTATAAAGGAACTAGTCCCGGCAAAACCCCGTAATTTGGAGCGAGAG	1569
Db	17300	TGCAGCGCAATCTATAAAGGAACTAGTCCCGGCAAAACCCCGTAATTTGGAGCGAGAG	17241
QY	1570	TGAGTGGGCGGGGACCCGACAGCGGAGCCGACCCCTTCTCTCCGGGCTCGGCGAGGCG	1629
Db	17240	TGAGTGGGCGGGGACCCGACAGCGGAGCCGACCCCTTCTCTCCGGGCTCGGCGAGGCG	17181
QY	1630	ASGGCGGGGAGCTC 1643	
Db	17180	AGGCGGGGAGCTC 17167	
RESULT 4			
AX092589			
LOCUS	AX092589	183999 bp	DNA linear PAT 21-MAR-2001
DEFINITION	Sequence 1 from Patent WO0115676.		
ACCESSION	AX092589		
VERSION	AX092589.1	GI:13444647	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.		
TITLE	Compositions and methods for modulating hdl cholesterol and triglyceride levels		
JOURNAL	Patent: WO 0115676-A 1 08-MAR-2001.		
FEATURES	University of British Columbia (CA) ; Xenon Genetics Inc. (CA) Location/Qualifiers		
source	1. .183999		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	49549 a 37944 c 41170 g 54950 t	386 others	
ORIGIN			
Query Match	91.2%;	Score 1498.6;	DB 6; Length 183999;
Best Local Similarity	97.9%;	Pred. No. 0;	
Matches 1614;	Conservative	2; Mismatches	10; Indels 23; Gaps
QY	16	GGCTCCACATGCACCTCCAGGCGCTGCTGG-CTCTCTTAATGGGTCTGCTCAGTGTT	73
Db	27234	GGCTCCACATGCACCTCCAGGCGCTGCTGGCGCTCTCTATGGCTCTGCTCCTCAGTGTT	27293
QY	74	GATAGAACCACTGATGTGAGTACCTTGGGCTTTGAG-CTGGGCTGGAGATCCTGTTGACTG	132
Db	27294	GATAGAACCACTGATGTGAGTACCTTGGGCTTTGAGCGCTGGAGATCCTGTTGACTG	27353
QY	133	TAGCATGAGGGGGCTTGT-CAGCTGAATCTCTGTATGCAGGCTGGTGGGAGTTCTGGAAT	191
Db	27354	TAGCATGAGGGGGCTTGTGAGCTGAATCTCTGTATGCAGGCTGGTGGGAGTTCTGGAAT	27413
QY	192	ATGATGGAGCTGAGGTTGGGAAGAACTAGGCTTTGGGAGCTCTCTCATGCGCACCTCA	251
Db	27414	ATGATGGAGCTGAGGTTGGGAAGAACTAGGCTTTGGGAGCTCTCTCATGCGCACCTCA	27473
QY	252	TCTGGCCCAAACTCAGTCCAAC-TGGAAGAGTCTAAATGTGAATCTGCCCTTCAAGGT	311
Db	27474	TCTGGCCCAAACTCAGTCCAAC-TGGAAGAGTCTAAATGTGAATCTGCCCTTCAAGGT	27533
QY	312	GGCTACAAAGTATCTTTCTCAAGTAGGACCTTTGGGCTCCACGTGCAATTCAGG	371
Db	27534	GGCTACAAAGTATCTTTCTCAAGTAGGACCTTTGGGCTCCACGTGCAATTCAGG	27593

Db	2838	CCTCTGGCTGGTCGACCGCAATCTATATAAGGAACATAGTCCCGGCAAAACCCGTAA	2897	repeat_region
		/rpt_family="Alu"		
		6625..6877		
Qy	1557	TGGGAGGAGAGTGAGTGGGGCGGGACCCGACGAGCGGACCGACTTCTCTCCCGG	1616	repeat_region
Db	2898	TGGGAGGAGAGTGAGTGGGGCGGGACCCGACGAGCGGACCGACTTCTCTCCCGG	2957	repeat_region
Qy	1617	GCTGGCGAGGCGAGGCGGGGAGGTC	1643	repeat_region
Db	2958	GCTGGCGAGGCGAGGCGGGGAGGTC	2984	repeat_region
RESULT 6				
AF275948		149034 bp	linear	PRI 17-JUL-2000
LOCUS				
DEFINITION		Homo sapiens ABCA1 (ABCA1) gene, complete cds.		
ACCESSION		AF275948		
VERSION		AF275948.1		
KEYWORDS		GI:9247085		
SOURCE				
ORGANISM		Homo sapiens.		
REFERENCE				
AUTHORS		Ekaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 149034); Santamarina-Jojo,S., Peterson,K., Knapper,C., Oiu,Y., Freeman,L., Cheng,J.F., Osorio,J., Remaley,A., Yang,X.P., Haudenschild,C., Prades,C., Chimini,G., Blackmon,E., Francois,T., Duverger,N., Rubin,E.N., Rosier,M., Deneffe,P., Fredrickson,D.S. and Brewer,H.B. Jr.		
TITLE		Complete genomic sequence of the human ABCA1 gene: analysis of the human and mouse ATP-binding cassette A promoter		
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7587-7592 (2000)		
MEDLINE		20345099		
PUBMED		10884428		
REFERENCE				
AUTHORS		2 (bases 1 to 149034); Santamarina-Jojo,S., Peterson,K.M., Knapper,C.L., Freeman,L.A., Remaley,A.T., Yang,X.P., Haudenschild,C.C., Blackmon,E.E., Francois,T.L. and Brewer,H.B. Jr.		
TITLE		Direct Submission		
JOURNAL		Submitted (08-JUN-2000) Molecular Disease Branch, National Institutes of Health, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA		
FEATURES				
source		1..149034		
gene		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		1454..148034		
		/gene="ABCA1"		
mRNA		join(1454..1674,25831..25989,40385..40478,45012..45153,46423..46541,67729..67850,70831..71007,83960..84052,89010..89250,91962..92101,92433..92549,96758..96955,97702..97907,98428..98604,100391..100613,102360..102581,103642..103846,104951..105064,106862..107033,108023..108154,109460..109602,109806..109943,110646..110866,112124..112196,113183..113385,115053..115191,115297..115410,116806..116954,118604..118728,119964..120062,123094..123283,124804..124898,126208..126240,127363..127468,128943..129017,129539..129708,130939..131116,133114..133229,133441..133485,134526..134649,135736..135865,136130..136250,137038..137100,138008..138114,140469..140610,140982..141116,142060..142163,142646..142738,143397..143640,144581..148034)		
		/gene="ABCA1"		
		/product="ABCA1"		
		join(1454..1674,25831..25923)		
		/gene="ABCA1"		
		2406..2678		
repeat_region		/rpt_family="Alu"		
repeat_region		4240..4509		
		/rpt_family="Alu"		
repeat_region		4998..5270		
		/rpt_family="Alu"		

TFVLELFDNKLNNINDILKSVLIFPHFCILGRGLIDMKVKNQAMADALERGENREYS
PLSWDLGNLFAMAVEGWFFLLITVLIORFFIRPRVNAKLSPLEADEDSDVRRERQ
RILDGGONDILEIKELTKYRRKRKPAVDRIQVPPGECFGLIGVNGAKSSFTFKM
LTGDTVTGDAFNKXNLSILNHEVHNMGYCPOFDAITELLTGREHVEFFALLRGV
PEKVKYGEWAKRLGLVKYGEKVAGNYSGRKRLSTAMALIGPPVVFIDEPTTG
MDPKARFLWNCALSVYKEGSRVLTSHSMECEALCTMAIWNNGRFRCLGSOVHLK
NRFGGYITVIAAGSNDDLKPVQDFGLAPGVSXKXKRNMLQYQLPSS:SSLARI
PSLSQKRRHIEDYSQPTLQDFVFNFAKQSDSDHLKLSLHKNTQVVDVAVLI
SPLODEKRESVY
29916..30180
/rpt_family="Alu"
repeat_region
31350..31354
/rpt_family="Alu"
repeat_region
31692..31973
/rpt_family="Alu"
repeat_region
36167..36440
/rpt_family="Alu"
repeat_region
37617..37892
/rpt_family="Alu"
repeat_region
37999..38166
/rpt_family="Alu"
repeat_region
38172..38269
/rpt_family="Alu"
repeat_region
38995..39271
/rpt_family="Alu"
repeat_region
41547..42585
/gene="ABCA1"
STS
/note="A009X28"
repeat_region
41593..41835
/rpt_family="Alu"
STS
41557..41679
/gene="ABCA1"
/note="A009X28"
repeat_region
42320..42584
/rpt_family="Alu"
repeat_region
44349..44626
/rpt_family="Alu"
repeat_region
48473..48669
/rpt_family="Alu"
repeat_region
50547..50614
/rpt_family="Alu"
repeat_region
54285..54418
/rpt_family="Alu"
repeat_region
54849..54988
/rpt_family="Alu"
repeat_region
55029..55226
/rpt_family="Alu"
repeat_region
55297..55344
/rpt_family="Alu"
repeat_region
58501..58935
/note="JTR"
repeat_region
61246..61489
/rpt_family="Alu"
repeat_region
62812..63077
/rpt_family="Alu"
repeat_region
64082..64306
/rpt_family="Alu"
repeat_region
67123..67382
/rpt_family="Alu"
repeat_region
68499..68731
/rpt_family="Alu"
repeat_region
69481..69760
/rpt_family="Alu"
repeat_region
71996..72233
/rpt_family="Alu"
repeat_region
72244..72501

Query Match 86.3%; Score 1418; DB 9; Length 149034;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 5; Indels 21; Gaps 8;
QY 113 CTTGGAGATCCTGTTGACGTAGCATGGAGGGGGCTTGT-CAGCTGAATGCTGTGATGCA 171

Db 1 CTTGGAGATCCTGTTGACGTAGCATGGAGGGGGCTTGTGACGTGAATGCTGTGATGCA 60
QY 172 GGTGGTGGAGTCTGGAAATATGATGGACTGGAGGAGAGAGAGTAGCTTGGGGC 231
Db 61 GGTGGTGGAGTCTGGAATATGATGGAGTGGAGTGGAGAGAGTAGCTTGGGGC 120
QY 232 AGTCTCTCATGCACCTCATTCIGGCCAAAACCTCAGGTCAAACITGGAAGAGTCTAAAT 291
Db 121 AGTCTCTCATGCACCTCATTCIGGCCAAAACCTCAGGTCAAACITGGAAGAGTCTAAAT 180
QY 292 GTGAATCTGCCCTTCAAGGTGGCTACAAAGGATATCTTTGTCAAGGTAGGAGACCTTGTGG 351
Db 181 GTGAATCTGCCCTTCAAGGTGGCTACAAAGGATATCTTTGTCAAGGTAGGAGACCTTGTGG 240
QY 352 CTTCCAGGTGCACCTCCAGGGGCTGCTT-GGCTCTCTACGGGTCTGTCTCAGTCTTC 410
Db 241 CTTCCAGGTGCACCTCCAGGGGCTGCTTGGGGCTCTCTACGGGTCTGTCTCAGTCTTC 300
QY 411 TATGAATCT-CCCTTCAGGGCAGATTCATATTTAGACTCTTCACASTTTCACTGAGTTT 459
Db 301 TATGAATCTCTCTTCAGGGCAGATTCATATTTAGACTCTTCACASTTTCACTGAGTTT 360
QY 470 TGGCCAGAAATAAGGTGACATTTAGTTTGGTGGCTTGAATGACTTAAATATATAGACA 529
Db 361 TGGCCAGAAATAAGGTGACATTTAGTTTGGTGGCTTGAATGACTTAAATATATAGACA 420
QY 530 TATGGTGTGTAGGCTCGATTCCTACTCTGCTTTTGTGCTTTTGTGCTTCTCCAGTCTTGG 589
Db 421 --TGGTGTGTAGGCTCGATTCCTACTCTGCTTTTGTGCTTTTGTGCTTCTCCAGTCTTGG 478
QY 590 GTAGTTTGTGCTCCCTTACAGCCAAAGCAACAGATAAGTTGGAGTCTGGAGTGGCTA 649
Db 479 GTAGTTTGTGCT-CCCTTACAGCCAAAGCAACAGATAAGTTGGAGTCTGGAGTGGCTA 537
QY 650 CATAATTTTACAGACTGCAATTTCTGGCTGCACTTCACAAATGTATACAAATATAATA 709
Db 538 CATAATTTTACAGACTGCAATTTCTGGCTGCACTTCACAAATGTATACAAATATAATA 597
QY 710 CAAGTCTGTGCTTTTATCAGAGGAGGCTGATCAATAATATGAATAATAAAGGGGCTG 769
Db 598 CAAGTCTGTGCTTTTATCAGAGGAGGCTGATCAATAATATGAATAATAAAGGGGCTG 657
QY 770 GTCCCATATTCCTGHTGTTTTTGT-----TTGTTTGTCTTTTGTGTTTGT 820
Db 658 GT-CCATATTTGCTGTTTTTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTGT 716
QY 821 GGCTCTCTCTCAATTTATGAAGAGAGAGAGAGAGAGAGTCTCTCTCGGGTCTCTGA 880
Db 717 GGCTCTCTCTCAATTTATGAAGAGAGAGAGAGAGAGAGTCTCTCTCGGGTCTCTGA 776
QY 881 GGGACCTGGGAGCTCAGGCTGGGAATCTCCAAGCAGTAGTCTCCCTATCAAAATCAA 940
Db 777 GGGACCTGGGAGCTCAGGCTGGGAATCTCCAAGCAGTAGTCTCCCTATCAAAATCAA 836
QY 941 AGTCAGGTTTCTGGGGGAAAACAAAAGCAGCCCAATACCCAGAGGACTGTCCGCTTC 1000
Db 837 AGTCAGGTTTCTGGGGGAAAACAAAAGCAGCCCAATACCCAGAGGACTGTCCGCTTC 896
QY 1001 CCTCAGCCCTCAGCTAGGCTTTGAAAGGAAAACAAAAGCAGCAAAATGATGGGCTC 1060
Db 897 CCTCAGCCCTCAGCTAGGCTTTGAAAGGAAAACAAAAGCAGCAAAATGATGGGCTC 956
QY 1061 CTGAGGAGATTCAGGCTAGAGCTCTCTCTCCCAATATCCCTCCCTCCGGGTGAGAAAC 1120
Db 957 CTGAGGAGATTCAGGCTAGAGCTCTCTCTCCCAATATCCCTCCCTCCGGGTGAGAAAC 1016
QY 1121 TAACAAAGGAAAATAAAATTCGGGAAAGCAGGATTTAGAGGAGCAAAATTCACCTGGTGC 1180
Db 1017 TAACAAAGGAAAATAAAATTCGGGAAAGCAGGATTTAGAGGAGCAAAATTCACCTGGTGC 1076
QY 1181 CTTTGGCTCCGGGACGTGGAGTCTGGGGCAGGCTAGAGTCTGGGGCAGGCTAGAGTCTTC 1240

Db	1077	CCTTGGCTGCCGGAAACGTGGAGTAGAGAGTCTGGGCCAGCCCGAGCCAGCCGCTTC	1136
Qy	1241	CCGCGCTCTTAGCGCCGCGGGCCGGCGGGAAGGGGAGCAGACCGCGGACCTTAA	1300
Db	1137	CGCGCGTCTTAGCGCGCGCGGGCGGGGAAGGGGACGACACCGGAGCCCTTAA	1196
Qy	1301	GACACCTGCTGACCCCTCCAC-----CCCCACCCACCCACCTCCGCCCACTCCCTAGA	1355
Db	1197	GACACCTGCTGACCCCTCCACCCACCCACCCACCCACCTCCGCCCACTCCCTAGA	1256
Qy	1356	TGTGTGTTGGGCGGTGAACGTGCGCCGTTTAAAGGGGGGGCCCGGGTCCACGTGCTTT	1415
Db	1257	TGTGTGTTGGGCGGTGAACGTGCGCCGTTTAAAGGGGGGGCCCGGGTCCACGTGCTTT	1316
Qy	1416	CTGTGTGAGTGAAGTGAACATTAACAGAGGCGGGAGCGGGCGGGAGGAGGAGC	1475
Db	1317	CTGTGTGAGTGAAGTGAACATTAACAGAGGCGGGAGGCGGGAGGAGGAGGAGC	1376
Qy	1476	ACAGGCTTTACCGATAGTAAGCTCTGCGCTCGTGCAGCCGAACTATAAAGGAACCTA	1535
Db	1377	ACAGGCTTTACCGATAGTAAGCTCTGCGCTCGTGCAGCCGAACTATAAAGGAACCTA	1436
Qy	1536	GTCCGGCAAAACCCCGTAATTCGAGCGAGAGTGAAGTGGGCGGGAGCCGAGAGCC	1595
Db	1437	GTCCGGCAAAACCCCGTAATTCGAGCGAGAGTGAAGTGGGCGGGAGCCGAGAGCC	1496
Qy	1596	GAGCCAGCCCTCTCTCCGGGCTCGGCAGGCGAGGCGGGGAGCTC	1643
Db	1497	GAGCCAGCCCTCTCTCTCCGGGCTCGGCAGGCGAGGCGGGGAGCTC	1544

RESULT 7

AX351031

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.

Regulatory nucleic acid sequences of the abcl gene

Patent: WO 0183746-A 3 08-NOV-2001;

Avantis Pharma S.A. (FR)

Location/Qualifiers

1..2893

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1524; Conservative

0; Mismatches

5; Indels

26; Gaps

10;

Qy

16

GCTCCACATGCACCTCCAGGCGCTGGTGG--CTCTTCTATGGGTGTGCTGAGTGT

73

Db

1345

CCCTCCACATGCACCTCCAGGCGCTGGTGGCGCTTCTATGCTCTGCTCCTGAGTGT

1404

Qy

74

GATAGAACACTGATGTGAGTACCTGGGCTTGAG--CGTGGCTCGAGATCCTGTGACTG

132

Db

1405

GATAGAACACTGATGTGAGTACCTGGGCTTGAGCGCTGGCTGGAGATCCTGTGACTG

1464

Qy

133

TACATGGAGCGGCTTGT--CAGCTGAATCTGTATCAGGTGCTGGAGTCTGGAAAT

191

Db

1465

TACATGGAGCGGCTTGTGTCAGCTGAATGCTGCA--CAGGTGTCGAGTCTGGAAAT

1524

Qy

192

ATCATGGAGCTGAGTGGGAAGATAGAGTCTGGGCGACCTCTCTCATGCCACCTCA

251

CDS

QY 1322 -----CCCCACCCACCCACCTCCCCCCAACCTCCCTAGATGCTGCGTGGCGGCTGAACG 1376
 |||||
 Db 2658 CCCCACCCACCCACCCACCTCCCCCCAACCTCCCTAGATGCTGCGTGGCGGCTGAACG 2717
 |||||
 QY 1377 TCGCCGCTTTAAGGGGGGGGGCGGCGGCTCCACGTCGCTTCGCTGAGTGAAGTGAACATACA 1436
 |||||
 Db 2718 TCGCCGCTTTAAGGGGGGGGGCGGCGGCTCCACGTCGCTTCGCTGAGTGAAGTGAACATACA 2777
 |||||
 QY 1437 TAAACAGAGCCCGGAACGGGGCGGAGGAGGAGGAGCAGAGCCTTGACCCATAGTAA 1496
 |||||
 Db 2778 TAAACAGAGCCCGGAAGGGGGCGGAGGAGGAGGAGCAGAGCCTTGACCCATAGTAA 2837
 |||||
 QY 1497 CCGTCGCGTCGTCGAGCCGCAATCTATATAAGGAATAGTCCCGGCAAAACCCC 1552
 |||||
 Db 2838 CCGTCGCGTCGTCGAGCCGCAATCTATATAAGGAATAGTCCCGGCAAAACCCC 2893
 |||||

RESULT 8

AF287262 201144 bp DNA linear PRI 29-APR-2001
 LOCUS Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
 DEFINITION and SNAP protein genes, complete cds.

ACCESSION AF287262
 VERSION AF287262.1 GI:13876612

KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
 TITLE Human and mouse ABCA1 comparative sequencing and transgenesis
 studies revealing novel regulatory sequences
 JOURNAL Genomics 73 (1), 66-76 (2001)
 MEDLINE 21251004
 PUBMED 11352567

REFERENCE
 AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUL-2000) Genome Science Department, Lawrence
 Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
 CA 94720, USA

FEATURES

source
 1..201144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 33931..181457
 /gene="ABCA1"
 mRNA
 join(33931..34151,58319..58476,72893..72886,77520..77661,
 78931..79049,101048..101169,104152..104328,117287..117379,
 122329..122569,125285..125424,125754..125870,
 130079..130276,131023..131228,131749..131925,
 133708..133930,135677..135898,136959..137163,
 138268..138381,140179..140350,141340..141471,
 142777..142919,143123..143260,143963..144183,
 145441..145513,146657..146859,148527..148575,
 148771..148884,150280..150428,152078..152202,
 153438..153536,156568..156757,158278..158372,
 159682..159714,160837..160942,162417..162491,
 166815..166959,168503..168628,169715..169844,
 170109..170229,171017..171079,171987..172093,
 174448..174589,174961..175095,176039..176142,
 176625..176717,177376..177619,178560..181457)
 /gene="ABCA1"
 exon
 33931..34151
 /product="ATP-binding cassette 1 sub-family A member 1"
 /gene="ABCA1"
 /number=1
 58319..58476
 /gene="ABCA1"
 /number=2

join(58411..58476,72893..72886,77520..77661,78931..79049,
 101048..101169,104152..104328,117287..117379,
 122329..122569,125285..125424,125754..125870,
 130079..130276,131023..131228,131749..131925,
 133708..133930,135677..135898,136959..137163,
 138268..138381,140179..140350,141340..141471,
 142777..142919,143123..143260,143963..144183,
 145441..145513,146657..146859,148527..148575,
 148771..148884,150280..150428,152078..152202,
 153438..153536,156568..156757,158278..158372,
 159682..159714,160837..160942,162417..162491,
 166815..166959,168503..168628,169715..169844,
 170109..170229,171017..171079,171987..172093,
 174448..174589,174961..175095,176039..176142,
 176625..176717,177376..177619,178560..178700)
 /gene="ABCA1"
 /codon_start=1
 /product="ATP-binding cassette 1 sub-family A member 1"
 /protein_id="AAK43526.1"
 /db_xref="GI:13876613"
 /translation="MACWPOLRLHKNLFRROTQCOLLELVANLELFLILSVRL
 SYPIEHECEFPNKAMPAGTLPWVGITICNANNPCFRPTGEPAGVGVNKNKSLV
 ARLPDSARLLYSQKDSIKMKRKLRTLQIKKSSNKLQDLVDNDFSGEYLH
 NLSIPKSTVDKMLRADVILHKVFLQGLVQLTSLCNSKSEEMIQDQVSELCGLP
 KEKLAARVLRNMDILKPIRLTNSPSPKSELAETKILLSLGTSLAQELFMK
 SWSDROEVMLTNVASSSTCIYOAVSRVCGHPGGGKIKSLNWNEDNYKALF
 GGNGTEAETFYDNTPTPCNDLMKNLESSPLRIWKALKLLYKLLYTPDPAT
 ROVNAENKTFQELAFPHLEGMWEEELSPKIFWFMENSQEMDLVRLMLLSDNDHWE
 QQLQDLMTAQDIAFLAKHPEDVQSSGSIYIWEAFNEQTAIRITISNFMCCNLN
 KLEP IATEVWLINKSMELLDERKFWAGIVFTGIPGSEILPHVHKIRKMDIDNERT
 NKIKDGVDPDPRADPEMDYVGMGFPAYLQDVVEQAIIRLVTKETKTDVMOQMPY
 PCYVDDIFLRVMSBPLEMTLAWISYVAILGIIVKEARLEKRLWRINGLDSILM
 /SWPISSELLPLLYSAGLVTILKGNLLPYSDPSVVFVLSFAVTVILQCLISTIF
 SRNLAAACGGIITFLYFLICVQWQDVGFTLKFASLLSPYAFGFCGCFALFE
 EFGIGVQMDNLFESPEEDGFNLTSVSMMLFNTFLYGVMTWYLEAVFPQYGIPIRW
 YPGCTKSYWFEESDEKSHPNOKRISEICEEPEHLKLGVSILNLYKVRDGMKV
 AVDGLANFYEGQITSLGHNGAGTKTMSILTLGTPPTSGTAYILGDIRSEKSTIR
 ONLQVCPQHNVLFDMLTVERHIFWYARLKLKSKHKAEMOMALDVLPSLKSXKI
 SOLSGQRKLSVALAGVSKSVLDEPTAGVDPYSRGWELLLKVRGRTLIILST
 HMDKEDVLDGRIALISHGKCCGCSLFLKQLQCTGYILTLVKKDVSSLCRNS
 STVYLKEDVSQSSDAGSDHESDILIDVSAISNLIRKEVSEARVEDIGHLEL
 TVYLPYPAKEGAEVELFHEIDRLSDLGISYISSETLEIEILKVAESGVDKQYK
 DCTPARRNRFAFGDKSCLRPFTEDDAADPNDSIDPESRETLLSGMDGKGYQVK
 GWKLTQOQFVALLIKRLIARRSKGFFAQLVPAFVCIALVSLVFPFGKYPSE
 LQPMYNEQVTVSNDAPEDGTLELLNLTDPGFCMCEGNPIPTPCQAGEEW
 TTAVPQTIMDLFQNGWTKMNPACOCSSDKIKMLPVCPPGAGGLPPOKQYV
 D-LQDLTGRIISDLTKYIVQIIAASKLKKIWENEFRYGGSGISGVNTOALPPSOEYV
 DAIKOMKHLKLAKDSADRLNSLGRMTGLDTRNNKVKWFKNKGWHAISFLNVIN
 NAILRANQKGNPSHYGITAFNEPZNLTKOOLSEVALMTSDVLSVICIFAMSFV
 PASFVFLIQBRVSKAKHLQTSIGVKPYIYLSNFMVDMCNVVPAILVILFICFQO
 KSYVSTNLVYLAALLLLYKSIPLMYPASVFKIPSTAVVLTSLNLFITINGSV
 TVFLLEFTDNKLNINDILKSVFLIFPHCLGRGLIDMWKQAMADALERGENRFS
 PLSWDLVGRNLFMAVEGVFELTVLYOYREFIPRPVNAKLSPLNGVDCAGKSTFKM
 ILTGDITVTRGDALFNKNSLGNIRHEHQMNGYCPQDAITELLTGRESVEFALLRGV
 PEKVGKGEWAIARKLGVKGYKAGNYSGNKKRKLSTAKMIGGPPVFLDEFTTG
 NDPKARFLWNCALSVKSGRSVLTSHSMECEALCTMAIMVNGRPGCLSGVHLK
 NRPFGGYTLVYVIRAGSNPDLPQDFGLAPGVSVLKKEHNNILOYQLPSSLSSLARI
 SFILQSKKRLHIEDIYSVSQTTLDOQVFNFAKQDQDDHLDKLSLEKQTVVDVAULT
 SFLQDEKVKESYV"
 72893..72986
 /gene="ABCA1"
 /number=3
 77520..77661
 /gene="ABCA1"
 /number=4
 78931..79049
 /gene="ABCA1"
 /number=5
 101048..101169
 /gene="ABCA1"
 /number=6

exon

exon

exon

exon

Query Match	75.9%	Score 1247.6	DB 9	Length 201144
Best Local Similarity	88.9%	Pred. No. 0		
Matches 1464	Conservative	0	Mismatches 129	Indels 54
Gaps				
QY	16	GGCTCCACATGCATCTCCAGGGCGCTGCTGG--CGCTTCATATGGGTCTGTCTCTGAGTGT	73	
DB	32410	GGCTCCACATGCATCTCCAGGGCGCTGCTGGCGCTCTCTTAAGCTGTCTGTCTGAGTGT	32469	
QY	74	GATAGAACCACTGATGTGAGTACCTGGCTTCAG--CGTGGCTCGAGATCTGTGACTG	132	
DB	32470	GATAGAACCACTGATGTGAGTACCTGGCTTCAGCGGTGGCTGGAGATCTGTGACTG	32529	
QY	133	TAGCATGAGGGGGCTTCT--CAGCTGAATCTGTATACAGGTGTGGAGTCTTGGAT	191	
DB	32530	TAGCATGAGGGGGCTTCTGTCAGCTGAATCTGTGATCAGGTGTGGAGTCTTGGAT	32589	
QY	152	ATGATGGAGCTGGAGGTGGGAAGAAAGTAGCTTGGGCGACTCTCTCATGCCACCTCA	251	
DB	32590	ATGATGGAGCTGGAGGTGGGAAGAAAGTAGCTTGGGCGACTCTCTCATGCCACCTCA	32649	
QY	252	TTCTGGCCAAAPACTCAGGTCAAACTGTGAAGAGICTAAATGTGAATCTGCCCTTCAAGT	311	
DB	32650	TTCTGGCCAAAPACTCAGGTCAAACTGTGAAGAGICTAAATGTGAATCTGCCCTTCAAGT	32709	
QY	312	GGCTACAAAGGTATCTTTGCTCAAGGTAGGAGACCTTGTGGCGTCCACGTGCACCTCCAG	371	
DB	32710	GGCTACAAAGGTATCTTTGCTCAAGGTAGGAGACCTTGTGGCGTCCACGTGCACCTCCAG	32769	
QY	372	GCTGTCT--GGCTCTTCTACGGCTCTGCTCTGAGTCTCTATGAATCT--CCCTTCAGGG	429	
DB	32770	GCTGTCTGGGGCTCTCTACGGGTCTGCTCTGAGTCTCTATGAATCTCTGAGTCTTCAGGG	32829	
QY	430	CAGATTCATATTTAGACTCTTTCACAGTTTGACCTGAGTTTGGCCCGAATAAGGTGCAT	489	
DB	32830	CAGATTCATATTTAGACTCTTTCACAGGGGATCAATAGGTTCCGCCATGATGG--	32883	
QY	490	TTAGTTTGTGTGCTTGATGAATGACTTAATATTTAGACATATGGTGTGTAGGCTGCAT	549	
DB	32884	-----ACTCAGTNNAGNCCCGCTCTCCNNNNNNNN	32916	
QY	550	TCTACCTCTTGGCTTTTGTGGCCCTCAGTGTCTTGGGTAGTTTGGTCCCGCTACA	609	
DB	32917	NNNNNNNNNNCTTTTGTGGCCCTCAGTGTCTTGGGTAGTTTGGTCCCGCTACA	32975	
QY	610	GCBAAGGCAACAGATAAGTTGGAGTCTGGAGTGGCTACATAATTTACACGACTGCA	669	
DB	32976	GCBAAGGCAACAGAGAAGTTGGAGTCTGGAGTGGCTACATAATTTACACGACTGCA	33035	
QY	670	ATTCCTGGGTGCACCTTCACAAATGTTATACAACTAAATACAGTCTGTGTTTATCA	729	
DB	33036	ATTCCTGGGTGCACCTTCACAAATGTTATACAACTAAATACAGTCTGTGTTTATCA	33095	
QY	730	CAGGAGGCTGATCAATATATGAATTAAGAGGGGCTGGCCATATGTTCTGTGTGT	789	
DB	33096	CAGGAGGCTGATCAATATATGAATTAAGAGGGGCTGGT--CCATATGTTCTGTGT	33154	
QY	790	TTTGTGTGTTGTCTTTTGTGTTTGT--GGCTCTCTCTCTCAATTTA	841	
DB	33155	TTTGTGTGTTGTCTTTTGTGTTTGT--GGCTCTCTCTCTCAATTTA	33214	
QY	842	TGAAGAGACGATGAAGTGTCTCTCGGTCTCTGAGGAGCTGGGAGCTCAGCT	901	
DB	33215	TGAGGAGAACCACTGAGGTTCTCTTTGGCCCTTAAAGAGCTAGTACCTTAAGCT	33274	
QY	902	GGGAATCTCCAGGCGATGAGTGGCTTACAAAATCAAAGTCCAGGTTGTGGGGGAA	961	
DB	33275	CGGGAATTTCCAGGCGTAATTCCTCTATATAAAAAAACACGAGGTGTGATCACAAA	33334	
QY	962	ACAAAGAGAGCCATTACCCAGAGGACTGTCCGCTCTCCCTCCACCCAGCTAGGCT	1021	
DB	33335	ACAAAGAGAGCCATTACCCAGAGGACTGTCTGCCCTTCCCTCAGCCAGGCTAGGCT	33394	


```
* 42936 45448: contig of 2513 bp in length
* 45449 45548: gap of 100 bp
* 45549 48116: contig of 2568 bp in length
* 48117 48216: gap of 100 bp
* 48217 52618: contig of 4402 bp in length
* 52619 52718: gap of 100 bp
* 52719 55592: contig of 3874 bp in length
* 55593 56692: gap of 100 bp
* 56693 59635: contig of 2943 bp in length
* 59636 59735: gap of 100 bp
* 59736 63661: contig of 3925 bp in length
* 63662 63761: gap of 100 bp
* 63762 68437: contig of 4676 bp in length
* 68438 68537: gap of 100 bp
* 68538 71458: contig of 2921 bp in length
* 71459 71558: gap of 100 bp
* 71559 76888: contig of 5330 bp in length
* 76889 76988: gap of 100 bp
* 76989 82113: contig of 5125 bp in length
* 82114 82213: gap of 100 bp
* 82214 88220: contig of 5007 bp in length
* 88221 88320: gap of 100 bp
* 88321 93498: contig of 5179 bp in length
* 93500 93599: gap of 100 bp
* 93600 97901: contig of 4302 bp in length
* 97902 98001: gap of 100 bp
* 98002 103016: contig of 5015 bp in length
* 103017 103116: gap of 100 bp
* 103117 109178: contig of 6062 bp in length
* 109179 109278: gap of 100 bp
* 109279 117307: contig of 8029 bp in length
* 117308 117407: gap of 100 bp
* 117408 124079: contig of 6672 bp in length
* 124080 124179: gap of 100 bp
* 124180 131281: contig of 7102 bp in length
* 131282 131381: gap of 100 bp
* 131382 138059: contig of 6678 bp in length
* 138060 138159: gap of 100 bp
* 138160 145491: contig of 7332 bp in length
* 145492 145591: gap of 100 bp
* 145592 157391: contig of 11800 bp in length
* 157392 157491: gap of 100 bp
* 157492 175064: contig of 17573 bp in length.
FEATURES
    source
        1..175064
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="R211-1M10"
            /clone_lib="R211-11 Human Male BAC"
        1..1003
            /note="assembly_fragment"
        1104..2634
            /note="assembly_fragment"
        2735..4415
            /note="assembly_fragment"
        4516..5795
            /note="assembly_fragment"
        5886..7879
            /note="assembly_fragment"
        7980..9686
            /note="assembly_fragment"
        9787..12253
            /note="assembly_fragment"
        12354..15228
            /note="assembly_fragment"
        15329..17200
            /note="assembly_fragment"
        17301..20131
            /note="assembly_fragment"
        20232..22567
            /note="assembly_fragment"
        22688..25707
            /note="assembly_fragment"
```

```
misc_feature 25808..28184
    /note="assembly_fragment"
misc_feature 28285..31338
    /note="assembly_fragment"
misc_feature 31439..34299
    /note="assembly_fragment"
misc_feature 34400..38318
    /note="assembly_fragment"
misc_feature 38419..42835
    /note="assembly_fragment"
misc_feature 42936..45448
    /note="assembly_fragment"
misc_feature 45549..48116
    /note="assembly_fragment"
misc_feature 48217..52618
    /note="assembly_fragment"
misc_feature 52719..55592
    /note="assembly_fragment"
misc_feature 55693..59635
    /note="assembly_fragment"
misc_feature 59736..63661
    /note="assembly_fragment"
misc_feature 63762..68437
    /note="assembly_fragment"
misc_feature 68538..71458
    /note="assembly_fragment
    vector_side:left"
misc_feature 71559..76888

Query Match 61.3%: Score 1007; DB 2; Length 175064;
Best Local Similarity 99.9%: Pred. No. 2e-257;
Matches 1018; Conservative 0; Mismatches 0; Indexes 1; Gaps 1;

Qy 625 ATAAAGTTGGAGGCTGGAGTGGCTACATAATTTACAGCACTGCAATCTCTGCGTGCAC 684
Db 2735 ATAAAGTTGGAGGCTGGAGTGGCTACATAATTTACAGCACTGCAATCTCTGCGTGCAC 2794

Qy 685 TTCACAAATGTATACAACTAAATACAAAGTCTCTGTTTATCACAGGAGGCTGATCA 744
Db 2795 TTCACAAATGTATACAACTAAATACAAAGTCTCTGTTTATCACAGGAGGCTGATCA 2854

Qy 745 ATATAATGAAATTAATAAGGGGCTGGTCCCATATGTTCTCTGTTTGTGTTGTTT 804
Db 2855 ATATAATGAAATTAATAAGGGGCTGGT-CCAATATGTTCTCTGTTTGTGTTGTTT 2913

Qy 805 CTITTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 864
Db 2914 CTITTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2973

Qy 865 CTCTCGGTCCTCTGAGGAGCTGGGAGCTCAGGCTGGGAACTCTCCAGGAGGCTGATGC 924
Db 2974 CTCTCGGTCCTCTGAGGAGCTGGGAGCTGGGAGCTGGGAACTCTCCAGGAGGCTGATGC 3033

Qy 925 GCCTATCAAAATCAAAAGTCCAGTTTCTGGGGGAAACAAACAGCCCATTTACCCAG 984
Db 3034 GCCTATCAAAATCAAAAGTCCAGTTTCTGGGGGAAACAAACAGCCCATTTACCCAG 3093

Qy 985 AGGACTGTCGGCTTCCCTTCACCCAGCTAGGCTTTGAAAGGAAACAAAGCAAGA 1044
Db 3094 AGGACTGTCGGCTTCCCTTCACCCAGCTAGGCTTTGAAAGGAAACAAAGCAAGA 3153

Qy 1045 CAAATGATTGGCTCTCTGAGGAGATTGAGCTTAGAGCTCTCTCTCCCCCAATCCCTCC 1104
Db 3154 CAAATGATTGGCTCTCTGAGGAGATTGAGCTTAGAGCTCTCTCTCCCCCAATCCCTCC 3213

Qy 1105 CTCGGCTGAGGAAACTAACAAGGAAAAAATAATTCGGGAAAGCAAGTTAGAGGAG 1164
Db 3214 CTCGGCTGAGGAAACTAACAAGGAAAAAATAATTCGGGAAAGCAAGTTAGAGGAG 3273

Qy 1165 CAAATTCCTACTGTCCTTCCCTTGGCTCCCGGAGACGTGACTAGAGCTCTGCGGCGGAGCC 1224
Db 3274 CAAATTCCTACTGTCCTTCCCTTGGCTCCCGGAGACGTGACTAGAGCTCTGCGGCGGAGCC 3333
```


[illegible][illegible]

DeArellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
 Garayna, S., Grant, C., Hagos, B., Heeford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, J., O'Donnell, P., Olivari, T. M., Peterson, K.,
 Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A., and Zody, N.

TITLE

JOURNAL

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705871.

All repeats were identified using RepeatMasker:

Smit, A. F. A. & Greer, P. (1996-1997)

<http://ftp.genome.washington.edu/JM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W13R

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N_13

 * NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 872 971: gap of 100 bp
 972 1834: contig of 863 bp in length
 1835 1934: gap of 100 bp
 1935 2804: contig of 870 bp in length
 2805 2904: gap of 100 bp
 2905 3745: contig of 841 bp in length
 3746 3845: gap of 100 bp
 3846 4696: contig of 851 bp in length
 4697 4796: gap of 100 bp
 4797 5640: contig of 844 bp in length
 5641 5740: gap of 100 bp
 5741 6540: contig of 800 bp in length
 6541 6640: gap of 100 bp
 6641 7503: contig of 869 bp in length
 7510 7603: gap of 100 bp
 7610 8479: contig of 870 bp in length
 8480 8579: gap of 100 bp
 8580 9430: contig of 851 bp in length
 9431 9530: gap of 100 bp
 9531 10376: contig of 846 bp in length
 10377 10476: gap of 100 bp
 10477 11322: contig of 846 bp in length
 11323 11422: gap of 100 bp
 11423 12302: contig of 880 bp in length
 12303 12402: gap of 100 bp
 12403 13280: contig of 878 bp in length
 13281 13390: gap of 100 bp
 13381 14241: contig of 861 bp in length
 14242 14341: gap of 100 bp
 14342 15196: contig of 855 bp in length
 15197 15296: gap of 100 bp
 15297 16123: contig of 827 bp in length

16124 16223: gap of 100 bp
 16224 17072: contig of 849 bp in length
 17073 17172: gap of 100 bp
 17173 18044: contig of 869 bp in length
 18045 18144: gap of 100 bp
 18145 19003: contig of 868 bp in length
 19010 19109: gap of 100 bp
 19110 19966: contig of 857 bp in length
 19967 20066: gap of 100 bp
 20067 20921: contig of 855 bp in length
 20922 21021: gap of 100 bp
 21022 21865: contig of 844 bp in length
 21866 21965: gap of 100 bp
 21966 22832: contig of 867 bp in length
 22833 22932: gap of 100 bp
 22933 23780: contig of 848 bp in length
 23781 23880: gap of 100 bp
 23881 24733: contig of 853 bp in length
 24734 24834: gap of 100 bp
 24834 25670: contig of 837 bp in length
 25671 25770: gap of 100 bp
 25771 26621: contig of 851 bp in length
 26622 26721: gap of 100 bp
 26722 27576: contig of 855 bp in length
 27577 27676: gap of 100 bp
 27677 28332: contig of 856 bp in length
 28333 28632: gap of 100 bp
 28633 29492: contig of 860 bp in length
 29493 29592: gap of 100 bp
 29593 30455: contig of 863 bp in length
 30456 30555: gap of 100 bp
 30556 31410: contig of 855 bp in length
 31411 31510: gap of 100 bp
 31511 32368: contig of 858 bp in length
 32369 32468: gap of 100 bp
 32469 33312: contig of 844 bp in length
 33313 33412: gap of 100 bp
 33413 34268: contig of 856 bp in length
 34269 34368: gap of 100 bp
 34369 35204: contig of 836 bp in length
 35205 35304: gap of 100 bp
 35305 36156: contig of 852 bp in length
 36157 36256: gap of 100 bp
 36257 37128: contig of 872 bp in length
 37129 37228: gap of 100 bp
 37229 38083: contig of 855 bp in length
 38084 38183: gap of 100 bp
 38184 39031: contig of 848 bp in length
 39032 39131: gap of 100 bp
 39132 40006: contig of 875 bp in length
 40007 40106: gap of 100 bp
 40107 40967: contig of 861 bp in length
 40968 41067: gap of 100 bp
 41068 41913: contig of 846 bp in length
 41914 42013: gap of 100 bp
 42014 42824: contig of 811 bp in length
 42825 42924: gap of 100 bp
 42925 43775: contig of 852 bp in length
 43777 43876: gap of 100 bp
 43877 44752: contig of 876 bp in length
 44753 44852: gap of 100 bp
 44853 45724: contig of 872 bp in length
 45725 45824: gap of 100 bp
 45825 46643: contig of 819 bp in length
 46644 46743: gap of 100 bp
 46744 47599: contig of 856 bp in length
 47600 47699: gap of 100 bp
 47700 48551: contig of 852 bp in length
 48552 48651: gap of 100 bp
 48652 49485: contig of 834 bp in length
 49486 49585: gap of 100 bp
 49586 50440: contig of 855 bp in length
 50441 50540: gap of 100 bp


```
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3746 3845: gap of 100 bp
* 3846 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 869 bp in length
* 7510 7609: gap of 100 bp
* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
```

```
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 38031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 43967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41069 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47500 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55329: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp
```

Query Match

Best Local Similarity 16.5%; Score 271.6; DB 2; Length 69570;

Matches 289; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1322 CCCCCCCCCACCCACCTCCCGCCCAACTCCCTAGATGTCTGCGGGGGGTGAACGTGGCC 1381

||||| ||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 67642 CCCCCNNCCCCACCACCTCCCGCCCAANTATAGATGTCTGCGGGGGGTGAACGTGGCC 67583

* 31554 32452: contig of 899 bp in length
* 32552 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 838 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42498: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length

* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match 14.0%; Score 229.8; DB 2; Length 90698;
Best Local Similarity 66.0%; Pred. No. 7.5e-50;
Matches 279; Conservative 0; Mismatches 141; Indels 3; Gaps 3:

QY 1221 AGCCCGGAGCCAGCGCTTCGGCGCGTCTAGGCGCGCGCGCGCGGAGGGG 1280
|||||
DB 87863 AGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87804
QY 1281 ACGCAGACCGGGACCCCTAAGACACCTGCTGTACCTCCACCCACCCACCTCC 1340
|||
DB 87803 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87744
QY 1341 CCCCAACTCCCTAGATGTCTGTGGCGGCTGAAGTCCGCGTTTAAAGGGCGGGCCCC 1400
|||||
DB 87743 CCACCTCCCGACCTAGATGTCTGTGGCGGCTGAAGTCCGCGTTTAAAGGGCGGGCCCC 87685
QY 1401 GCCTCCAGTCTTCTCTGTGAGTGACTCACTACATACACAGAGCGGGACGGGGG 1460
|||||
DB 87684 GGCTCCAGTCTTCTCTGTGAGTGACTCACTACATACACAGAGCGGGACGGGGG 87626
QY 1461 GGGAGGAGGAGAGACAGCGGCTTTGACCGATAGTAACTCTCGCTCGGTGCGAGCGGAAT 1520
|||||
DB 87625 GGGAGAGGAGAGACAGCGGCTTTGACCGATAGTAACTCTCGCTCGGTGCGAGCGGAAT 87566
QY 1521 CTATAAAGGAAGTACTAGTCCCGGCAAAACCCCGTAAATTCGGAGCGAGAGTGAAGTGGGCC 1580
|||||
DB 87565 CTATAAAGGAAGTACTAGTCCCGGCAAAACCCCGTAAATTCGGAGCGAGAGTGAAGTGGGCC 87507
QY 1581 GGGAGCGGAGAGCGAGCGGCGGCTTCCTCCCGGCGTGGCGAGCGGCGGCGGAG 1640
|||||
DB 87506 GGGAGCGGAGAGCGGCGGCGGCTTCCTCCCGGCGTGGCGAGCGGCGGCGGAG 87447
QY 1641 CTC 1643
|||
DB 87445 CTC 87444

RESULT 15
AF258627
LOCUS
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
partial; cds.
ACCESSION AF258627
VERSION AF258627.1 GI:7769707
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of hABCI gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
FEATURES
Location/Qualifiers
source
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"


```
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABC1"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AA269513.1"
/db_xref="GI:7769708"
/translation="MACHPQLRLLLWKLUFRROTQOLLLELVAMPLEFLLILSVRL
STPPYEQHECHFPNAMPFAGLPPVQGLICNANNPCFRYPPTGAPGVVGNFNKS"
BASE COUNT 152 a 198 c 150 g 156 t 1 others
ORIGIN

Query Match 10.5%; Score 173; DB 9; Length 697;
Best Local Similarity 100.0%; Pred.No.5.6e-35;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1471 AGAGCACAGGCTTTGACCGATAGTACCTCTCGCTCGGTGCGCGCAATCTATAAAAGG 1530
Db 1 AGAGCACAGGCTTTGACCGATAGTACCTCTCGCTCGGTGCGCGCAATCTATAAAAGG 60

QY 1531 AACTAGTCCCGGCAGAAACCCCTAATTGGAGCGAGAGTGAAGTGGGGCGGACCCGCA 1590
Db 61 AACTAGTCCCGGCAGAAACCCCTAATTGGAGCGAGAGTGAAGTGGGGCGGACCCGCA 120

QY 1591 GAGCCGAGCGGACCCCTCTCTCCCGGGCTGCGGCAGGCGAGGCGGGGAGCTC 1643
Db 121 GAGCCGAGCGGACCCCTCTCTCCCGGGCTGCGGCAGGCGAGGCGGGGAGCTC 173
```

Search completed: February 4, 2003, 00:59:00
Job time : 5346.97 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 : Search time 397.962 seconds
(without alignments)
10165.006 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1394_1532
Perfect score: 139
Sequence: 1 gggcccggtccacgtgct.....agccgaatctataaaagaa 139

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054540 seqs, 14531402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_om.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_rnd.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	139	100.0	1167	9	HSA252201	AJ252201 Homo sapi
2	139	100.0	1643	6	AX060715	AX060715 Sequence
3	139	100.0	1643	6	AX060894	AX060894 Sequence
4	139	100.0	69570	2	AC021246	AC021246 Homo sapi
5	139	100.0	96717	9	AL359182	AL359182 Human DNA
6	139	100.0	175064	2	AC012230	AC012230 Homo sapi
7	139	100.0	183999	6	AX092589	AX092589 Sequence
8	137.4	98.8	1167	9	AF258623s1	AF258623 Homo sapi
9	137.4	98.8	2893	6	AX351031	AX351031 Sequence
10	137.4	98.8	3231	6	AX351029	AX351029 Sequence
11	137.4	98.8	149034	9	AF275948	AF275948 Homo sapi
12	137.4	98.8	201144	9	AF287262	AF287262 Homo sapi
13	126.2	90.8	69570	2	AC021246	AC021246 Homo sapi
14	113.2	81.4	90698	2	AC021345	AC021345 Homo sapi
15	106.2	76.4	186889	2	AL807243	AL807243 Mus muscu
16	106.2	76.4	278572	10	AF287263	AF287263 Mus muscu
17	104	74.8	145833	2	AC125837	AC125837 Rattus no
18	62	44.5	697	9	AF258627	AF258627 Homo sapi
19	36.4	26.2	38	6	AX092823	AX092823 Sequence
20	34.6	24.9	193159	9	AC006946	AC006946 Homo sapi
21	33.4	24.0	129090	2	AP003931	AP003931 Oryza sat
22	33.4	24.0	133902	2	AC102375	AC102375 Mus muscu
23	33.4	24.0	218091	2	AC121802	AC121802 Mus muscu
24	32.8	23.5	107553	9	BSJ136014	BSJ136014 Human DNA
25	32.6	23.5	235198	10	AL645571	AL645571 Mouse DNA
26	32.4	23.3	138151	2	AC094070	AC094070 Rattus no
27	32.4	23.3	150994	2	AC111413	AC111413 Rattus no
28	32.2	23.2	125066	9	AC074286	AC074286 Homo sapi
29	32.2	23.2	156785	2	AC018541	AC018541 Homo sapi
30	32.2	23.2	170924	2	AC087619	AC087619 Homo sapi
31	32.2	23.2	171813	2	AC022722	AC022722 Homo sapi
32	32.2	23.2	185218	2	AC126911	AC126911 Rattus no
33	32.2	23.2	187994	2	AC126146	AC126146 Rattus no
34	32.2	23.2	192139	9	AC021955	AC021955 Homo sapi
35	32.2	23.2	215936	9	CNS05TCA	AL355076 Human chr
36	32	23.0	187359	2	AC068886	AC068886 Homo sapi
37	32	23.0	188425	2	AC036554	AC024854 Homo sapi
38	32	23.0	214229	9	AC011899	AC011899 Homo sapi
39	31.8	22.9	64523	2	AC100987	AC100987 Mus muscu
40	31.8	22.9	164290	2	AC096454	AC096454 Rattus no
41	31.6	22.7	132596	2	AC013046	AC013046 Drosophil
42	31.6	22.7	169101	3	AC008317	AC008317 Drosophil
43	31.6	22.7	212185	9	AC024651	AC024651 Homo sapi
44	31.6	22.7	329362	3	AE003681	AE003681 Drosophil
45	31.4	22.6	215889	2	AC092196	AC092196 Canis fam

ALIGNMENTS

RESULT 1
HSA252201
LOCUS
DEFINITION Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, 5'UTR and promoter region.
ACCESSION AJ252201
VERSION AJ252201.1 GI:12053757
KEYWORDS ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1167)
Porsch-Ozcurmez,M., Langmann,T., Heimerl,S., Borsukova,H., Kaminski,W.E., Drobni,W., Honer,C., Schumacher,C. and Schmitz,G.

Anderson, S., Baldwin, J., Baroz, N., Beckerly, R., Beda, F.,
 Boguslavsky, L., Bouhgalter, B., Brown, A., Burkett, G., Castle, A.,
 Choepell, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
 DeArrellanc, K., Dewar, K., Domino, M., Doyle, M., Fenesor, J.,
 Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Haqos, B., Heaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kane, L., Karatas, A., Klein, J.,
 Landers, T., Lehoczy, J., Levine, R., Lie, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tirelli, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A., and Zody, M.

TITLE JOURNAL

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705871.

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W13R

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N_10

* NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 871: contig of 871 bp in length
 872 971: gap of 100 bp
 972 1834: contig of 863 bp in length
 1835 1934: gap of 100 bp
 1935 2804: contig of 870 bp in length
 2805 2904: gap of 100 bp
 2905 3745: contig of 841 bp in length
 3746 3845: gap of 100 bp
 3846 4696: contig of 851 bp in length
 4697 4796: gap of 100 bp
 4797 5640: contig of 844 bp in length
 5641 5740: gap of 100 bp
 5741 6540: contig of 800 bp in length
 6541 6640: gap of 100 bp
 6641 7509: contig of 869 bp in length
 7510 7609: gap of 100 bp
 7610 8479: contig of 870 bp in length
 8480 8579: gap of 100 bp
 8580 9430: contig of 851 bp in length
 9431 9530: gap of 100 bp
 9531 10376: contig of 846 bp in length
 10377 10476: gap of 100 bp
 10477 11322: contig of 846 bp in length
 11323 11422: gap of 100 bp
 11423 12302: contig of 880 bp in length
 12303 12402: gap of 100 bp
 12403 13280: contig of 878 bp in length
 13281 13380: gap of 100 bp
 13381 14241: contig of 861 bp in length
 14242 14341: gap of 100 bp

14342 15196: contig of 855 bp in length
 15197 15296: gap of 100 bp
 15297 16123: contig of 827 bp in length
 16124 16223: gap of 100 bp
 16224 17072: contig of 849 bp in length
 17073 17172: gap of 100 bp
 17173 18041: contig of 869 bp in length
 18042 18141: gap of 100 bp
 18142 19009: contig of 868 bp in length
 19010 19109: gap of 100 bp
 19110 19966: contig of 857 bp in length
 19967 20066: gap of 100 bp
 20067 20921: contig of 855 bp in length
 20922 21021: gap of 100 bp
 21022 21865: contig of 844 bp in length
 21866 21965: gap of 100 bp
 21966 22832: contig of 867 bp in length
 22833 22932: gap of 100 bp
 22933 23780: contig of 848 bp in length
 23781 23880: gap of 100 bp
 23881 24733: contig of 853 bp in length
 24734 24833: gap of 100 bp
 24834 25670: contig of 837 bp in length
 25671 25770: gap of 100 bp
 25771 26621: contig of 851 bp in length
 26622 26721: gap of 100 bp
 26722 27576: contig of 855 bp in length
 27577 28532: contig of 856 bp in length
 28533 28632: gap of 100 bp
 28633 29492: contig of 860 bp in length
 29493 29592: gap of 100 bp
 29593 30455: contig of 863 bp in length
 30456 31410: contig of 855 bp in length
 31411 31510: gap of 100 bp
 31511 32368: contig of 858 bp in length
 32369 32468: gap of 100 bp
 32469 33312: contig of 844 bp in length
 33313 33412: gap of 100 bp
 33413 34268: contig of 856 bp in length
 34269 34368: gap of 100 bp
 34369 35204: contig of 836 bp in length
 35205 35304: gap of 100 bp
 35305 36156: contig of 852 bp in length
 36157 36256: gap of 100 bp
 36257 37128: contig of 872 bp in length
 37129 37228: gap of 100 bp
 37229 38083: contig of 855 bp in length
 38084 38183: gap of 100 bp
 38184 39031: contig of 848 bp in length
 39032 39131: gap of 100 bp
 39132 40006: contig of 875 bp in length
 40007 40106: gap of 100 bp
 40107 40967: contig of 861 bp in length
 40968 41067: gap of 100 bp
 41068 41913: contig of 846 bp in length
 41914 42013: gap of 100 bp
 42014 42824: contig of 811 bp in length
 42825 42924: gap of 100 bp
 42925 43776: contig of 852 bp in length
 43777 43876: gap of 100 bp
 43877 44752: contig of 876 bp in length
 44753 44852: gap of 100 bp
 44853 45724: contig of 872 bp in length
 45725 45824: gap of 100 bp
 45825 46643: contig of 819 bp in length
 46644 46743: gap of 100 bp
 46744 47599: contig of 856 bp in length
 47600 47699: gap of 100 bp
 47700 48551: contig of 852 bp in length
 48552 48651: gap of 100 bp
 48652 49485: contig of 834 bp in length


```

QY 121 GCCGAATCTATAAAGGAA 139
|||||
DQ 17296 GCCGAATCTATAAAGGAA 17278

RESULT 6
AC012230 Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
LOCUS pieces
DEFINITION AC012230
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1M10
Unpublished
2 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelac,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Gallagan,J., Gardyna,S., Grant,S., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Liedt,C., Locke,K., MacDonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severi,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirell,A., Vassiliev,E., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-Oct-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2510
Center clone name: L_M10
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version: C.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-contigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1004 1103: contig of 1003 bp in length
* 1004 1103: gap of 100 bp

```

* 157382 157491: gap of 100 bp
 * 157492 175064: contig of 17573 bp in length.

FEATURES

source
 1. .175064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-1M10"
 /clone_lib="RPC1-11 Human Male BAC"
 1. .1003
 /note="assembly_fragment"
 1104. .2634
 /note="assembly_fragment"
 2735. .4415
 /note="assembly_fragment"
 4516. .5785
 /note="assembly_fragment"
 5886. .7579
 /note="assembly_fragment"
 7980. .9586
 /note="assembly_fragment"
 9787. .12253
 /note="assembly_fragment"
 12354. .15228
 /note="assembly_fragment"
 15329. .17200
 /note="assembly_fragment"
 17301. .20131
 /note="assembly_fragment"
 20232. .22587
 /note="assembly_fragment"
 22688. .25707
 /note="assembly_fragment"
 25808. .28184
 /note="assembly_fragment"
 28285. .31338
 /note="assembly_fragment"
 31439. .34299
 /note="assembly_fragment"
 34400. .38318
 /note="assembly_fragment"
 38419. .42835
 /note="assembly_fragment"
 42936. .45448
 /note="assembly_fragment"
 45549. .48116
 /note="assembly_fragment"
 48217. .52618
 /note="assembly_fragment"
 52719. .56592
 /note="assembly_fragment"
 56693. .59635
 /note="assembly_fragment"
 59736. .63661
 /note="assembly_fragment"
 63762. .68437
 /note="assembly_fragment"
 68538. .71458
 /note="assembly_fragment"
 clone_end.SP6
 vector_side:left"
 71559. .76888

Query Match 100.0% Score 139; DB 2; Length 175064;
 Best Local Similarity 100.0%; Pred. No. 6.8e-32;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCCCGGCTCCACGTCGTTCTGCTGAGTGAACATACATAAAGAGCGCGGAA 60
 Db 3503 GGGCCCGGCTCCACGTCGTTCTGCTGAGTGAACATACATAAAGAGCGCGGAA 3562
 QY 61 CGGGCGGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 120
 Db 3563 CGGGCGGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 3622

QY 121 GCCGAATCTATAAAGGAA 139
 Db 3623 GCCGAATCTATAAAGGAA 364;
 RESULT 7
 LOCUS AX092589 18399 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 1 from Patent WO0115676.
 ACCESSION AX092589
 VERSION AX092589.1 GI:13444647
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 18399)
 AUTHORS Hayden,M.R., Brooks-Wilson,A.R., Plimstone,S.N. and Clee,S.M.
 TITLE Compositions and methods for modulating hdi cholesterol and triglyceride levels
 JOURNAL Patent: WO 0115676-A 1 08-MAR-2001;
 FEATURES Location/Qualifiers
 source
 1.183999
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 49549 a 37844 c 41170 g 54950 t 386 others
 ORIGIN
 Query Match 100.0% Score 139; DB 6; Length 183999;
 Best Local Similarity 100.0%; Pred. No. 6.8e-32;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCCCGGCTCCACGTCGTTCTGCTGAGTGAACATACATAAAGAGCGCGGAA 60
 Db 28631 GGGCCCGGCTCCACGTCGTTCTGCTGAGTGAACATACATAAAGAGCGCGGAA 28650
 QY 61 CGGGCGGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 120
 Db 28691 CGGGCGGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 28750
 QY 121 GCCGAATCTATAAAGGAA 139
 Db 28751 GCCGAATCTATAAAGGAA 28769
 RESULT 8
 LOCUS AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
 DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
 promoter and exon 1.
 ACCESSION AF258623
 VERSION AF258623.2 GI:8677405
 KEYWORDS
 SEGMENTS
 1 of 4
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1167)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
 TITLE Analysis of hABCL gene 5' end: additional peptide sequence,
 promoter region, and four polymorphisms
 JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
 REFERENCE 2 (bases 224 to 1167)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
 TITLE Direct Submission
 JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA

```

3 (bases 1 to 1167;
Pallinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aulizet,B.E., Fielding,C.J. and Kane,J.P.
Direct Submission
Submitted (23-JUN-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
Sequence update by submitter
COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.
FEATURES
source
    location/Qualifiers
    1..1167
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="9"
    /map="9c31"
    /name="ABCA1"
    224..864
    /name="ABCA1"
    845..1147
    /name="ABCA1"
    /number=1
BASE COUNT 272 a 313 c 328 g 248 t
ORIGIN
Query Match 98.8%; Score 137.4; DB 9; Length 1167;
Best Local Similarity 99.3%; Pred. No. 3.7e-31;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 60
Db 768 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 827
QY 61 CGGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 120
Db 828 GGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 887
QY 121 GCCGAATCTATAAAGGAA 139
Db 888 GCCGAATCTATAAAGGAA 906
RESULT 9
AX351031
LOCUS AX351031 2893 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 3 from Patent WO0183746.
ACCESSION AX351031
VERSION AX351031.1 GI:18616387
KEYWORDS human.
SOURCE
    ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
    Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P.,
    Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
    Regulatory nucleic acid sequences of the abcl gene
    Patent: WO 0183746-A 3 08-NOV-2001;
    Aventis Pharma S.A. (FR)
FEATURES
    source
    1..2893
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT 745 a 681 c 769 g 698 t
ORIGIN
Query Match 98.8%; Score 137.4; DB 6; Length 2893;
Best Local Similarity 99.3%; Pred. No. 3.3e-31;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 60
Db 2735 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 2794
QY 61 CGGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 120

```

```

Db 2795 GGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 2854
QY 121 GCCGAATCTATAAAGGAA 139
Db 2855 GCCGAATCTATAAAGGAA 2873
RESULT 10
AX351029
LOCUS AX351029 3231 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 1 from Patent WO0183746.
ACCESSION AX351029
VERSION AX351029.1 GI:18616385
KEYWORDS human.
SOURCE
    ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
    Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P.,
    Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
    Regulatory nucleic acid sequences of the abcl gene
    Patent: WO 0183746-A 1 08-NOV-2001;
    Aventis Pharma S.A. (FR)
FEATURES
    source
    1..3231
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT 809 a 773 c 876 g 773 t
ORIGIN
Query Match 98.8%; Score 137.4; DB 6; Length 3231;
Best Local Similarity 99.3%; Pred. No. 3.3e-31;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 60
Db 2735 GGGCCCGGGTCCAGTCTTCCTGCTGAGTGACTGAACACTACATAAAGAGCGCGGAA 2794
QY 61 CGGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 120
Db 2795 GGGCCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTGCA 2834
QY 121 GCCGAATCTATAAAGGAA 139
Db 2855 GCCGAATCTATAAAGGAA 2873
RESULT 11
AX275948
LOCUS AX275948 149034 bp DNA linear PRI 17-JUL-2000
DEFINITION Homo sapiens ABCA1 (ABCA1) gene, complete cds.
ACCESSION AP275948
VERSION AP275948.1 GI:9247085
KEYWORDS
    source
    ORGANISM Homo sapiens.
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
    (bases 1 to 149034)
    Santamarina-Fojo,S., Peterson,K., Knapper,C., Qiu,Y., Freeman,L.,
    Cheng,J.F., Osorio,J., Remaley,A., Yang,X.P., Haudenschild,C.,
    Prades,C., Chimini,G., Blackmon,E., Francois,I., Duverger,N.,
    Rubin,E.M., Rosier,M., Deneffe,P., Fredrickson,D.S. and Brewer,H.B.
    Jr.
    Complete genomic sequence of the human ABCA1 gene: analysis of the
    human and mouse ATP-binding cassette A promoter
    Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
    JOURNAL 20345099
    MEDLINE 10884628
    PUBMED 10884628
REFERENCE 2
    (bases 1 to 149034)
    Santamarina-Fojo,S., Peterson,K.M., Knapper,C.L., Freeman,L.A.,

```

July 5th

repeat_region 44349..44626
/rpt_family="Alu"
repeat_region 48473..48669
/rpt_family="Alu"
repeat_region 50547..50514
/rpt_family="Alu"
repeat_region 54285..54418
/rpt_family="Alu"
repeat_region 54849..54988
/rpt_family="Alu"
repeat_region 55029..55226
/rpt_family="Alu"
repeat_region 55297..55344
/rpt_family="Alu"
repeat_region 58501..58935
/note="LIR"
repeat_region 61246..61489
/rpt_family="Alu"
repeat_region 62812..63077
/rpt_family="Alu"
repeat_region 64082..64306
/rpt_family="Alu"
repeat_region 67123..67382
/rpt_family="Alu"
repeat_region 68499..68731
/rpt_family="Alu"
repeat_region 69481..69760
/rpt_family="Alu"
repeat_region 71998..72233
/rpt_family="Alu"
repeat_region 72244..72501

Query Match 98.88; Score 137.4; DB 9; Length 149034;
Best Local Similarity 99.38; Pred. No. 2:1e-31;
Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCGGCTCCAGTCCTTCTGCTGAGTGACTACATAACAGAGCGCGGAA 60
Db 1295 GGGCCCGGCTCCAGTCCTTCTGCTGAGTGACTACATAACAGAGCGCGGAA 1354
QY 61 CGGGCGGGAGGAGGAGACAGAGCTTTGACCGATAGTAACCTCGCTCGGTGCA 120
Db 1355 GGGCGGGAGGAGGAGGAGACAGAGCTTTGACCGATAGTAACCTCGCTCGGTGCA 1414
QY 121 GCCGAATCTATAAAGCAA 139
Db 1415 GCCGAATCTATAAAGCAA 1433
RESULT 12
AF287262
LOCUS AF287262 201144 bp DNA linear PRI 29-APR-2001
DEFINITION Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
and SNAP protein genes, complete cds.
ACCESSION AF287262
VERSION AF287262.1 GI:13876612
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 201144)
Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
Human and mouse ABCA1 comparative sequencing and transgenesis
studies revealing novel regulatory sequences
Genomics 73 (1), 66-76 (2001)
JOURNAL MEDLINE 21251004
PUBMED 11352567
REFERENCE 2 (bases 1 to 201144)
Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.F.
AUTHORS Direct Submission
TITLE Submitted (13-JUL-2000) Genome Science Department, Lawrence

Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
CA 94720, USA
FEATURES
source 1..201144
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
33931..181457
/gene="ABCA1"
join(33931..34151,58319..58476,72893..72986,77520..77661,
78931..79049,101048..101169,104152..104328,117287..117379,
122329..122569,125285..125424,125754..125870,
130079..130276,131023..131228,131749..131925,
133708..133930,135677..135898,136959..137163,
138268..138381,140179..140350,141340..141471,
142777..142919,143123..143260,143663..144183,
145441..145513,146657..146859,148527..148575,
148771..148884,150280..150426,152078..152202,
153438..153536,156568..156757,158278..158372,
159682..159714,160837..160942,162417..162491,
163013..163182,164413..164590,166588..166703,
168815..168959,168505..168628,169715..169844,
170109..170229,171017..171079,171987..172093,
174448..174569,174961..175095,176039..176142,
176625..176717,177376..177619,178560..181457)
/gene="ABCA1"
/product="ATP-binding cassette 1 sub-family A member 1"
33931..34151
/gene="ABCA1"
/number=1
58319..58476
/gene="ABCA1"
/number=2
join(58411..58476,72893..72986,77520..77661,78931..79049,
101048..101169,104152..104328,117287..117379,
123329..122569,125285..125424,125754..125870,
130079..130276,131023..131228,131749..131925,
133708..133930,135677..135898,136959..137163,
138268..138381,140179..140350,141340..141471,
142777..142919,143123..143260,143663..144183,
145441..145513,146657..146859,148527..148575,
148771..148884,150280..150426,152078..152202,
153438..153536,156568..156757,158278..158372,
159682..159714,160837..160942,162417..162491,
163013..163182,164413..164590,166588..166703,
168815..168959,168505..168628,169715..169844,
170109..170229,171017..171079,171987..172093,
174448..174569,174961..175095,176039..176142,
176625..176717,177376..177619,178560..181457)
/gene="ABCA1"
/codon_start=1
/product="ATP-binding cassette 1 sub-family A member 1"
/protein_id="AAK4326.1"
/db_xref="GI:13876613"
/translation="MACPQRLLLKMLTFRRCQQLLELVANWPFLLIISVL
SYPPYOEHCFFPNKAMPSAGTLPWOGIICNANNPCRYPTGEAPVGVNKNISV
ARLFSDDARLLLYSKDTSMDKRVLTQIQKSSNKLQDFIVDNTFSFGLYH
NLSPKSTVDKMLRADVILHKVLOGVOLHLJSLCNGSKSEMIOLGQDEVSGLGLP
KELAAAEVRLSRNMDILKPLRLNSTSPSPSKELAEATKILHSLGTLAQLFSMR
SWSDROEVMLTNVNSSSSTQIQVSRIVCGHGGGLKLSLNWYEDNYKALF
GGNDEDAETFDYDNTTFCNDLMKNLESPLRTIWKALKPLVGLKILYTPDTPAT
RQNAEYNTKTFELAVHLEGMWELSPKWTFMENSOEMDLVRLDSDNDHFE
QLDGLDWT AQDIVAFIAKPEYSSNGSVYTWRAFNETNQAIRTSRMECVLN
KLPIATEVWLINKMELDERKFWAGIVTGTITPGSIELPHKVKYKIRMDIGNERT
NKIKDGVDPGPDADPEMDRYVGGFAYLDVVEQAIIVLTGKTKVYWOQPY
PCVDDIFLWMSRSMPLMTLAWISVAVILKGVYKEARKEKEMRINGLNSILW
FSWISSILPLVLSAGLLVILKGNLLPYSDPSVFLSVFAVITLOCFLISILF
SRANAAACGGIIFLYLPYLCVAMQDVGTLLKIFASLPVDFGCEGFALFE
EQGTGVQWDLNFSPEVEDGNLTISVSMMLNTFLYGVMTWYIEAVFPQYGPWP
YFPTKSYWFGESDEKSHGSKRISIEICMEEPHKLGLVSIQNLVKVYRGKKY
AVDGLAVFEGQITFELGNHAGKKTYSILGLPPTSGTAYILGKDIRSEMSTIR
QNLGVCPQHNVLFDMITVEEHIWFYARLGLSKRKHVKAEMQALDGLDPSLKSRT
SOLSGGORKLSVALPVGSKYVILDETAGDVPYSRRGIWELLKRYGRTIILST
HHMDEADVLCRIATLISGKLCQVSSSLFLKNQIGTYLILVKKDVESSLSCRNSS

Gardyna, S., Grant, G., Hagos, B., Yeaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lebecky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Sartos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A., and Zody, M.

Direct Submission

TITLE

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL

On Jul 13, 2000 this sequence version replaced gi:6705871.

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N_10

* NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 872 971: contig of 871 bp in length
 872 971: gap of 100 bp
 1835 1934: contig of 863 bp in length
 1835 1934: gap of 100 bp
 1935 2804: contig of 870 bp in length
 2805 2904: gap of 100 bp
 2905 3745: contig of 841 bp in length
 3746 3845: gap of 100 bp
 3846 4596: contig of 851 bp in length
 4597 4796: gap of 100 bp
 4797 5540: contig of 844 bp in length
 5641 5740: gap of 100 bp
 5741 6540: contig of 800 bp in length
 6541 6640: gap of 100 bp
 6641 7509: contig of 869 bp in length
 7510 7609: gap of 100 bp
 7610 8479: contig of 870 bp in length
 8480 8579: gap of 100 bp
 8580 9430: contig of 851 bp in length
 9431 9530: gap of 100 bp
 9531 10376: contig of 846 bp in length
 10377 10476: gap of 100 bp
 10477 11322: contig of 846 bp in length
 11323 11422: gap of 100 bp
 11423 12302: contig of 880 bp in length
 12303 12402: gap of 100 bp
 12403 13280: contig of 878 bp in length
 13281 13380: gap of 100 bp
 13381 14241: contig of 861 bp in length
 14242 14341: gap of 100 bp
 14342 15196: contig of 855 bp in length
 15197 15295: gap of 100 bp
 15297 16123: contig of 827 bp in length
 16124 16223: gap of 100 bp
 16224 17072: contig of 849 bp in length
 17073 17172: gap of 100 bp
 17173 18041: contig of 869 bp in length
 18042 18141: gap of 100 bp
 18142 19009: contig of 868 bp in length
 19100 19109: gap of 100 bp
 19110 19966: contig of 857 bp in length
 19967 20066: gap of 100 bp
 20067 20921: contig of 855 bp in length
 20922 21021: gap of 100 bp
 21022 21865: contig of 844 bp in length
 21866 21965: gap of 100 bp
 21966 22832: contig of 867 bp in length
 22833 22932: gap of 100 bp
 22933 23780: contig of 848 bp in length
 23781 23880: gap of 100 bp
 23881 24733: contig of 853 bp in length
 24734 24833: gap of 100 bp
 24834 25670: contig of 837 bp in length
 25671 25770: gap of 100 bp
 25771 26621: contig of 851 bp in length
 26622 26721: gap of 100 bp
 26722 27576: contig of 855 bp in length
 27577 27676: gap of 100 bp
 27677 28532: contig of 856 bp in length
 28533 28632: gap of 100 bp
 28633 29432: contig of 860 bp in length
 29433 29592: gap of 100 bp
 29593 30455: contig of 863 bp in length
 30456 30555: gap of 100 bp
 30556 31410: contig of 855 bp in length
 31411 31510: gap of 100 bp
 31511 32358: contig of 858 bp in length
 32359 32468: gap of 100 bp
 32469 33312: contig of 844 bp in length
 33313 33412: gap of 100 bp
 33413 34268: contig of 856 bp in length
 34269 34368: gap of 100 bp
 34369 35204: contig of 836 bp in length
 35205 35304: gap of 100 bp
 35305 36156: contig of 852 bp in length
 36157 36256: gap of 100 bp
 36257 37128: contig of 872 bp in length
 37129 37228: gap of 100 bp
 37229 38083: contig of 855 bp in length
 38084 38183: gap of 100 bp
 38184 39031: contig of 848 bp in length
 39032 39131: gap of 100 bp
 39132 40006: contig of 875 bp in length
 40007 40106: gap of 100 bp
 40107 40967: contig of 861 bp in length
 40968 41067: gap of 100 bp
 41068 41913: contig of 846 bp in length
 41914 42013: gap of 100 bp
 42014 42824: contig of 811 bp in length
 42825 42924: gap of 100 bp
 42925 43776: contig of 852 bp in length
 43777 43876: gap of 100 bp
 43877 44752: contig of 876 bp in length
 44753 44852: gap of 100 bp
 44853 45724: contig of 872 bp in length
 45725 45824: gap of 100 bp
 45825 46643: contig of 815 bp in length
 46644 46743: gap of 100 bp
 46744 47599: contig of 856 bp in length
 47600 47699: gap of 100 bp
 47700 48551: contig of 852 bp in length
 48552 48651: gap of 100 bp
 48652 49485: contig of 834 bp in length
 49486 49585: gap of 100 bp
 49586 50440: contig of 855 bp in length
 50441 50540: gap of 100 bp
 50541 51404: contig of 864 bp in length
 51405 51504: gap of 100 bp


```
* 16778 17678: contig of 901 bp in length
* 17779 17778: gap of 100 bp
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18680 18779: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19633 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22579: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23668: gap of 100 bp
* 23669 24554: contig of 886 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length
* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
```

```
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match      81.4%; Score 113.2; DB 2; Length 90698;
Best Local Similarity 92.8%; Pred. No. 5e-24;
Matches 129; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 GGGCCCGGGCTCCACGCTTTTGTGTGAGTGACTGAACTACATAACACAGACGCCGGAA 60
    |||||
Db 87691 GGGCCCGGGCTCCACGCTTTTGTGTGAGTGACTGAACTACATAACACAGACGCCGGAA 87633
    |||||
QY 61 CGGGCGGGGAGGAGGAGACACAGCTTGACCGTAGTACCTCGCGCTCGGTGCA 120
    |||||
Db 87632 CGGGCGGGGAGGAGGAGACACAGCTTGACCGTAGTACCTCGCGCTCGGTGCA 87573
    |||||
QY 121 GCCGAATCTATAAAGGAA 139
    |||||
Db 87572 GCCGAATCTATAAAGGAA 87554
    |||||

RESULT 15
AL807243/C
LOCUS      186889 bp DNA linear HTG 30-JUN-2002
DEFINITION Mus musculus chromosome 4 clone RP23-25D17, *** SEQUENCING IN
PROGRESS ***; 24 unordered pieces.
ACCESSION AL807243
VERSION AL807243.5 GI:21568234
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
            McLay, K.
REFERENCE 1 (bases 1 to 186889);
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. Email enquiries:
tumquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:21668136.
----- Genome Center
```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humphries@sanger.ac.uk
----- Project Information
Center project name: BM25D17

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 175959 bases at least Q40
Consensus quality: 179463 bases at least Q30
Consensus quality: 182069 bases at least Q20
Insert size: 184589; sum-of-contigs
Insert size: 189466; 5.5% error; agarose-fp
Quality coverage: 4.72x in Q20 bases; sum-of-contigs Quality
coverage: 4.82x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 4849: contig of 4849 bp in length
4850 4949: gap of 100 bp
4950 22620: contig of 17671 bp in length
22621 22720: gap of 100 bp
22721 29963: contig of 7243 bp in length
29964 30053: gap of 100 bp
30054 40087: contig of 10024 bp in length
40088 40187: gap of 100 bp
40188 44438: contig of 4251 bp in length
44439 44538: gap of 100 bp
44539 48960: contig of 4422 bp in length
48961 49050: gap of 100 bp
49051 56007: contig of 6947 bp in length
56008 56107: gap of 100 bp
56108 58284: contig of 2177 bp in length
58285 58384: gap of 100 bp
58385 60786: contig of 2402 bp in length
60787 60886: gap of 100 bp
60887 66707: contig of 5821 bp in length
66708 66807: gap of 100 bp
66808 92102: contig of 25295 bp in length
92103 92202: gap of 100 bp
92203 99909: contig of 7707 bp in length
99910 100009: gap of 100 bp
100010 103760: contig of 3751 bp in length
103761 103860: gap of 100 bp
103861 110765: contig of 6905 bp in length
110766 110865: gap of 100 bp
110866 114685: contig of 3820 bp in length
114686 114785: gap of 100 bp
114786 125959: contig of 11174 bp in length
125960 126059: gap of 100 bp
126060 128749: contig of 2690 bp in length
128750 128849: gap of 100 bp
128850 132373: contig of 3524 bp in length
132374 132473: gap of 100 bp
132474 134930: contig of 2457 bp in length
134931 135030: gap of 100 bp
135031 139301: contig of 4271 bp in length
139302 139401: gap of 100 bp
139402 144448: contig of 5047 bp in length
144449 144548: gap of 100 bp
144549 147074: contig of 2526 bp in length
147075 147174: gap of 100 bp
147175 156317: contig of 9143 bp in length
156318 156417: gap of 100 bp
156418 186889: contig of 30472 bp in length.
```

FEATURES

```
Source
1. .186889
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
/clone_lib="RP23-25D17"
/clone_lib="RP23-25D17"
1. .4849
/notes="assembly_fragment:01920"
fragment_chain:1
4950. .22620
/notes="assembly_fragment:01761"
fragment_chain:1
22721. .29963
/notes="assembly_fragment:00140"
fragment_chain:1
30064. .40087
/notes="assembly_fragment:00335"
fragment_chain:1
40188. .44438
/notes="assembly_fragment:00241"
fragment_chain:1
44539. .48960
/notes="assembly_fragment:01430"
fragment_chain:1
49061. .56007
/notes="assembly_fragment:01667"
fragment_chain:1
56108. .58284
/notes="assembly_fragment:00965"
fragment_chain:2
58385. .60786
/notes="assembly_fragment:01660"
fragment_chain:2
60887. .66707
/notes="assembly_fragment:00374"
fragment_chain:2
66808. .92102
/notes="assembly_fragment:01792"
fragment_chain:2
92203. .99909
/notes="assembly_fragment:01082"
fragment_chain:2
100010. .103760
/notes="assembly_fragment:00358"
fragment_chain:2
103861. .110765
/notes="assembly_fragment:01182"
fragment_chain:3
110866. .114685
/notes="assembly_fragment:01582"
fragment_chain:3
114786. .125959
/notes="assembly_fragment:00043"
fragment_chain:3
126060. .128749
/notes="assembly_fragment:00352"
fragment_chain:3
128850. .132373
/notes="assembly_fragment:00217"
fragment_chain:4
132474. .134930
/notes="assembly_fragment:01737"
fragment_chain:4
135031. .139301
/notes="assembly_fragment:00487"
139402. .144448
/notes="assembly_fragment:00557"
144549. .147074
/notes="assembly_fragment:01342"
147175. .156317
/notes="assembly_fragment:01564.0"
156418. .186889
/notes="assembly_fragment:01773"
```

BASE COUNT 52900 a 39593 c 40710 g 51338 t 2348 others
ORIGIN
Query Match 76.4%; Score 106.2; DB 2; Length 185889;
Best Loca. Similarity 89.4%; Pred. No. 6.1e-22;
Matches 126; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
Qy 1 GGGCCCCGGCTCCACGTGCTTTCTGCTGAGTGACTGAACATACATAAACAGAGCGCGGAA 60
Db 99260 GGGCCAICTCTCCACGTGCTTTCTGCTGAGTGACTGAACATACATAAACAGAGCGCGGAA 99201
Qy 61 CGGGGGCGGG--AGGAGGGAGGACAGGCTTTGACCGATAGTACCTCTGGGCTCGGTG 118
Db 99200 GGGGGGGGGGAAAGAGGGAGAGACAGCGTTTGACCGGTAGTAAACCCCGCGCTCGGA 99141
Qy 119 CAGCCGAATCTATAAAGGAA 139
Db 99140 CAGCCGAATCTATAAAGGAA 99120

Search completed: February 4, 2003, 01:43:20
Job time : 748.962 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:39:45 ; Search time 3108.84 Seconds
(without alignments)
8559.21C Million cell updates/sec

Title: US-09-596-141C-3

Perfect score: 1643

Sequence: 1 gaattccttgctggtgctc.....cagggcagggcgggagctc 1643

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST : *

1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_estl: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	245.4	14.9	289	10	AW748338 RC5-BT025
2	111.4	6.8	467	9	AA527406
3	89.6	5.4	736	9	AA527406 ng37802.s
4	89.6	5.4	547	9	AU135588
5	89.6	5.4	763	9	AL698654 DRF26686N
6	75	4.6	292	14	AU121731
					Z44377 HSC125081

7	75	4.6	1004	11	BC034824	Homo sapi
8	60.8	3.7	1201	17	CNS015W3	Drosophila
9	60.2	3.7	998	12	BC678861	BG678861
10	54.8	3.3	298	17	CNS04M2X	Tetradon
11	53.8	3.3	844	17	CNS0052P	AL296898
12	53	3.2	237	9	AU222489	Drosophila
C 13	52.4	3.2	910	17	CNS006ON	AU222489
14	51.2	3.1	832	17	CNS015XC	Drosophila
C 15	51	3.1	681	17	CNS02FE9	AL105026
C 16	50.6	3.1	925	17	CNS0091P	AL194922
17	50.6	3.1	1101	17	CNS017SY	AL053613
C 18	50.4	3.1	844	17	CNS0052P	AL108460
C 19	50.4	3.1	870	17	CNS005EZ	Drosophila
C 20	48.4	2.9	1004	14	BQ944204	AL064371
C 21	48	2.9	900	17	AG081217	BQ944204
C 22	47.5	2.9	1009	17	CNS010EW	AG081217
C 23	47.6	2.9	1101	17	CNS01720	AL098882
C 24	47.4	2.9	1059	14	BQ680479	AL107514
C 25	47.4	2.9	1101	17	CNS017V2	BQ680479
C 26	47.2	2.9	477	9	AL513813	AL108536
C 27	47.2	2.9	925	17	CNS0091P	Drosophila
C 28	47.2	2.9	942	17	AG042878	AL053013
C 29	47	2.9	658	12	BF968907	AG042878
C 30	47	2.9	843	17	CNS00CS1	BF968907
C 31	47	2.9	973	17	CNS008OF	AL059666
C 32	47	2.9	997	17	CNS006GN	Drosophila
C 33	46.8	2.8	431	9	AL513947	AL052332
C 34	46.8	2.8	976	17	CNS0068V	Drosophila
C 35	46.6	2.8	978	17	AG030617	AL063041
C 36	46.6	2.8	1101	17	CNS00B3U	AG030617
C 37	46.2	2.8	527	9	AL514325	AL056389
C 38	46.2	2.8	875	17	AG043475	Drosophila
C 39	46.2	2.8	1136	14	BQ943816	AG043475
C 40	46.2	2.8	1417	14	BME10236	BQ943816
C 41	46	2.8	758	17	AG073529	BM10236
C 42	45.8	2.8	203	10	AW150511	AG073529
C 43	45.8	2.8	471	10	BE503803	AW150511
C 44	45.8	2.8	523	9	AL514015	BE503803
C 45	45.8	2.8	766	17	AG041031	AL514015

ALIGNMENTS

RESULT 1
AW748338/c 289 bp mRNA linear EST 28-APR-2000
LOCUS RC5-BT0252-271099-012-c10 BT0252 Homo sapiens cDNA, mRNA sequence.
DEFINITION RC5-BT0252-271099-012-c10 BT0252 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW748338
VERSION AW748338.1 GI:7663270
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 289)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed

Sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6&t2=RC6-BT0252-271099-012-c10&t3=1998-10-27&t4=1)

Seq primer: puc 18 forward
 High quality sequence start: 146
 High quality sequence stop: 289.
 Location/Qualifiers
 1. .289

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT0252"
 /dev_stage="Adult"

/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESIES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 94 a 70 c 51 g 74 t

ORIGIN

Query Match 14.9%; Score 245.4; DB 10; Length: 289;
 Best Local Similarity 96.9%; Pred. No. 6.3e-39;
 Matches 282; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 678 GCTGCACCTCACAATGATACAACTAAATACAGTCTCTGTTTATCACAGGGAGG 737
 Db 289 GCTGCACCTCACAATGATACAACTAAATACAGTCTCTGTTTATCACAGGGAGG 230
 QY 738 CTGATCAATATAATGAATATAAAGGGGCTGGTCCCATATTGCTGGTTTGTGTTG 797
 Db 229 CTGATCAATATAATGAATATAAAGGGGCTGGTCCCATATTGCTGGTTTGTGTTG 171
 QY 798 TTTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 857
 Db 170 TTTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 112
 QY 858 GATGTCCTCTCGGGTCTCTGAGGGACCTGGGAGCTCAGGCTGGGAATCTCCAAAGCA 917
 Db 111 GATGTCCTCTCGGGTCTCTGAGGGACCTGGGAGCTCAGGCTGGG-ATCTCCAAAGCA 53
 QY 918 GTAGTGCCTATCAAAATCAAACTCCAGTTCTGTTGGGGGAAACAA 968
 Db 52 GTAGTGCCTATCAAAATCAAACTCCAGTTCTGTTGGGGGAAACAA 2

RESULT 2
 AA527406
 LOCUS na37c02.s1 NCI_CGAP_C03 Homo sapiens cDNA clone IMAGE:936962 3', linear EST 21-AUG-1997
 DEFINITION mRNA sequence.
 ACCSSION AA527406
 VERSION AA527406.1 GI:2269475
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 467)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1057 Std Error: 0.00
 Seq primer: -40n13 fwd. ET from Amersham
 High quality sequence stop: 380.
 Location/Qualifiers
 1. .467

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NAG3:936962"
 /clone_lib="NCI_CGAP_C03"
 /sex="pooled"
 /tissue_type="colon"
 /lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization."

BASE COUNT 72 a 118 c 166 g 111 t

ORIGIN

Query Match 5.8%; Score 111.4; DB 9; Length 467;
 Best Local Similarity 99.1%; Pred. No. 3e-12;
 Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 816 TTGTGGGCTCTCTCTCAATTTATGAAGAGGAGCAGTAGATGTTCTCTCGGGTCC 875
 Db 1 TTTGTGGGCTCTCTCTCAATTTATGAAGAGGAGCAGTAGATGTTCTCTCGGGTCC 60
 QY 876 TCTCAGGACCTGGGGAGCTCAGGCTGGGAATCTCCAAAGGAGTAGTGCCT 928
 Db 61 TCTCAGGACCTGGGGAGCTCAGGCTGGGAATCTCCAAAGGAGTAGTGCCT 113

RESULT 3
 AC135588
 LOCUS PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA linear EST 02-AUG-2002
 DEFINITION sequence.
 AC135588
 VERSION AC135588.1 GI:10956127
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 736)
 Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.

HRI human cDNA project
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kiserazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986

Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

Location/Qualifiers
 1. .736

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE1002437"

```

/clone_lib="PLACE1"
/tissue_type="placenta"
/notes="Vector: pME18SFL3"
BASE COUNT 163 a 199 c 199 g 170 t 5 others
ORIGIN
Query Match 5.4%; Score 89.4; DB 9; Length 736;
Best Local Similarity 98.9%; Pred. No. 5.3e-08;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GTAAATGCGAGCAGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 1612
Dd 1 GTAGTGGCAGCAGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 60

QY 1613 CCGCGTCCGCGAGCGAGCGGCGGGAGCTC 1643
Dd 61 CCGGCTCGCGCAGCGCAGCGCGGGAGCTC 91

RESULT 4
AL698654 547 bp mRNA linear EST 21-MAR-2002
LOCUS DKFZp686N12109.5f.686 (synonym: hicc3) Homo sapiens cDNA clone
DEFINITION DKFZp686N12109 5', mRNA sequence.
ACCESSION AL698654
VERSION AL698654
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Cassenhuber,J. and Wiemann
,S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
MIPS
Am Klopferstutz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686N12109) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
location/Qualifiers
1..547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686N12109"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pTriplEx2; Site_1: SfiI; Site_2: SfiI;
cDNA collection"
BASE COUNT 112 a 154 c 157 g 124 t
ORIGIN
Query Match 5.4%; Score 89; DB 9; Length 547;
Best Local Similarity 100.0%; Pred. No. 8.5e-08;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1555 AATTGCGAGCGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 1614
Dd 2 AATTGCGAGCGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 61

QY 1615 GCGCTCGGCGAGCGGCGGGAGCTC 1643
Dd 62 GCGCTCGGCGAGCGGCGGGAGCTC 90

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pME18SFL3"
/tissue_type="mammary gland"
/notes="Vector: pME18SFL3"
BASE COUNT 137 a 205 c 260 g 158 t
ORIGIN
Query Match 5.4%; Score 89; DB 9; Length 763;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1555 AATTGCGAGCGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 1614
Dd 1 AATTGCGAGCGAGTGTAGTGGCGGGACCCGAGCGAGCGGACCCCTTCTCTC 60

QY 1615 GCGCTCGGCGAGCGGCGGGAGCTC 1643
Dd 61 GCGCTCGGCGAGCGGCGGGAGCTC 89

RESULT 6
AL698654 292 bp mRNA linear EST 14-NOV-1994
LOCUS HSC1ZB081 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION c-12b08, mRNA sequence.
ACCESSION 244377
VERSION 244377
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 292)
AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
Sebastiani-Kabaktchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
TITLE

```

JOURNAL
MEDLINE
COMMENT

and its expression
C. R. Acad. Sci. III, Sci. Vie 3:8 (2), 263-272 (1995)
95277534
Contact: Genethon
GenexPress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_idt: C; Genexpress_sequence_idt: ylc-1zb08
Seq primer: (-21)M13_universal.

FEATURES

source

```

1..292
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="c-1zb08"
  /clone_lib="normalized infant brain cDNA"
  /sex="Female"
  /tissue_type="total brain"
  /dev_stage="3 months old"
  /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex:female; dev_stage=3 months old; isolate-muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B. Soares, Psychiatry Dept. Columbia University, USA. Normalization method: Bento Soares, P.N.A.S in press"
BASE COUNT      50 a   87 c   96 g   56 t
ORIGIN

```

```

Query Match      4.6%; Score 75; DB 14; Length 292;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1569 GTGAGTGGGGCGGGACCCGAGAGCCGAGCCCTTCTCTCCGGGCTCGCGCAGGG 1628
|||||
Db 1 GTGAGTGGGGCGGGACCCGAGAGCCGAGCCCTTCTCTCCGGGCTCGCGCAGGG 60
|||||

Qy 1829 CAGGGCGGGGAGCTC 1643
|||||
Db 61 CAGGGCGGGGAGCTC 75
|||||

```

```

RESULT 7
BC034824
LOCUS      BC034824      1004 bp      mRNA      linear      HTC 26-JUL-2002
DEFINITION Homo sapiens, clone IMAGE:4749735, mRNA.
ACCESSION  BC034824
VERSION     BC034824.1 GI:21961568
KEYWORDS   HTC.
SOURCE      Homo sapiens.
ORGANISM   Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mamalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530,
USA
REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk
           Email: cgapbs-femail.nih.gov
           Tissue Procurement: James Cleaver, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Baylor College of Medicine Human Genome
           Sequencing Center
           Center code: BCM-HGSC

```

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Günaratne, P.H., Garcia, A.M., Lu, X., Rulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: iRAK Plate: 52 Row: d Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5915657
This clone has the following problem: incomplete processing.

FEATURES

source

```

1..1004
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:4749735"
  /tissue_type="Skin, squamous cell carcinoma"
  /clone_lib="NCI_CGAP_Skn4"
  /lab_host="DH10B"
  /note="Vector: pCMV-SPORT6.ccdB"
BASE COUNT      306 a   234 c   230 g   234 t
ORIGIN

```

```

Query Match      4.6%; Score 75; DB 11; Length 1004;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1569 GTGAGTGGGGCGGGACCCGAGAGCCGAGCCCTTCTCTCCGGGCTCGCGCAGGG 1628
|||||
Db 1 GTGAGTGGGGCGGGACCCGAGAGCCGAGCCCTTCTCTCCGGGCTCGCGCAGGG 60
|||||

```

```

Qy 1629 CAGGGCGGGGAGCTC 1643
|||||
Db 61 CAGGGCGGGGAGCTC 75
|||||

```

```

RESULT 8
CNS015W3
LOCUS      CNS015W3      1201 bp      DNA      linear      GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
           BACN15123 of DrosBAC library from Drosophila melanogaster (fruit
           fly), genomic survey sequence.
ACCESSION  AL105981
VERSION     AL105981.1 GI:5619397
KEYWORDS   GSS.
SOURCE      Drosophila melanogaster.
ORGANISM   Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL
COMMENT    Determination of this BAC-end sequence was carried out as part of a
           collaboration with the European Drosophila Genome Project (EDGP) -
           http://www.edgp.ebi.ac.uk - . This Drosophila melanogaster BAC
           library (Dros BAC) was made by Alain Billaud at CEPH (Centre
           d'Etude du Polymorphisme Humain) with funding provided by a MRC
           project grant. The DNA was prepared from embryos by Alain Bucheton
           and Genevieve Payan. It has been constructed in the vector
           pBelobAC11.

```

FEATURES

source

```

1..1201
Location/Qualifiers
  /organism="Drosophila melanogaster"
  /db_xref="taxon:7227"
  /clone="BACN15123"
  /clone_lib="DrosBAC"
  /p_asmid="pBelobAC11"
  /note="end : SP6"

```

```

BASE COUNT      255 a      276 c      278 g      235 t      157 others
ORIGIN

Query Match      3.7%; Score 60.8; DB 17; Length 1201;
Best Local Similarity 37.3%; Pred. No. 0.027;
Matches 109; Conservative 51; Mismatches 132; Indels 0; Gaps 0;

Qy 1115 GGAACATAACAGAAAGAAATATCGGGAAGACAGATTTAGAGAGCAATATCCAC 1174
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 857 KRAAMAAWAAAKTKMTGATATMKDTGGGMAAATAAGAAWAAWDDADTTKKK 916

Qy 1175 TGTGGCCCTGGCTCGCGGGAACGTGACTAGAGAGCTTCGCGCCAGCCCGAGCCAG 1234
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 917 TKTKGKCGSGCGSCCTSCSSCARASMGSGSGGGGCGCGCCCGCGSGSCNGC 976

Qy 1235 CGCTTCGCGCGCTTCTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGA 1294
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 977 CCGSSCCGSCGSCGSSSSGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGCGCC 1036

Qy 1295 CCCTAAGACACCTCTGTACCTCCACCCACCCACCCACCCACCCACCCCAACTCCCTAG 1354
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1037 CCGSCCCCCCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCGCGCC 1096

Qy 1355 AIGTGTCTGCGCGCGCTGAAGCTCGCGCGCTTTAAGGCGCGCGCGCGCGCGCGCC 1406
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1097 CCGSSSSSCGSGGGGSGSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGSC 1148

RESULT 9
BG678861      998 bp      mRNA      linear      EST 01-MAY-2001
LOCUS      602624760F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5',
DEFINITION      mRNA sequence.
ACCESSION      BG678861
VERSION      BG678861.1 GI:13910258
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE      1 (bases 1 to 998)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: rcapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10603 Row: g Column: 15
High quality sequence stop: 860.
Location/Qualifiers
      . . 998
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:4749735"
      /clone_lib="NCI_CGAP_Skn4"
      /tissue_type="squamous cell carcinoma"
      /lab_host="DH10B (T1 phage-resistant)"
      /note="Organ: skin; Vector: pCMV-SpO16; Site:1: NotI;
      Site:2: SalI; Cloned unidirectionally. Primer: oligo dT.
      Average insert size 1.5kb. Library constructed by Life
      Technologies. Note: this is a NCI-CGAP library."

BASE COUNT      285 a      233 c      244 g      236 t
ORIGIN

Query Match      3.7%; Score 60.2; DB 12; Length 998;
Best Local Similarity 59.0%; Pred. No. 0.038;
Matches 65; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1571 GAGTGGGCGCGGAGCCGAGAGCGAGCCGAGCCCTCTCTCCGCGCGCGGAGGCA 1630
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1 GTGAGTGGGCGGAGCCGAGAGCGGAGCCGAGCCCTCTCTCCGCGCGCGGAGGCA 60

Qy 1631 GGGCGGGGAGCTC 1643
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 61 GGGCGGGGAGCTC 73

RESULT 10
CNS04M2X      298 bp      DNA      linear      GSS 21-MAY-2000
LOCUS      Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION      119N17 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL296898
VERSION      AL296898.1 GI:8035478
KEYWORDS      GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 298)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bertot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
TITLES      Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 298)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bertot, A. and
Weissenbach, J.
TITLES      Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE      3 (bases 1 to 298)
AUTHORS      Genoscope.
TITLES      Direct Submission
JOURNAL      Submitted (12-APR-2000)
COMMENT      This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
      . . 298
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone="119N17"
      /clone_lib="G"
      /note="Genoscope sequence ID : C0BGL119CG09LP1-end : T7"

BASE COUNT      18 a      106 c      80 g      15 t      79 others
ORIGIN

Query Match      3.3%; Score 54.8; DB 17; Length 298;
Best Local Similarity 40.5%; Pred. No. 0.72;
Matches 81; Conservative 33; Mismatches 86; Indels 0; Gaps 0;

Qy 1180 CCCTTGGCTGCGGGAAGCTGACTAGAGTCTCGCGCGAGCCCGAGCCAGCGGTT 1239
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 82 CCNKKSCCCVDVCGCCVCVSGSGSGSCSCSCSSSSSSSSSSSSSSSSSSSSSSSS 141

Qy 1240 CCGCGCGCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 1299
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 142 GSGSSCCSCGSGSGSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCNATC 201

Qy 1300 AGACACCTGTCTACCTCCACCCACCCACCCACCCACCCACCCACCTAGATGTG 1359
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 202 NCNMCNCGCTCTCNMCNMCNMCNMCNMCNMCNMCNMCNMCNMCNMCNMCNMCN 261

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CorpuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:46:48 ; Search time 448.725 Seconds
(without alignments)
8245.650 Million cell updates/sec

Title: US-09-596-141C-3

Perfect score: 1643

Sequence: 1 gaattcctgctggtgctc.....caggcgagggcgaggagctc 1643

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1643	100.0	1643	22 AAF24681	Nucleotide sequenc
2	1643	100.0	1643	22 AAF24703	Nucleotide sequenc
3	1498.6	91.2	183999	22 AAF2831	Human ABC1 genomic
4	1485.4	90.4	3231	24 AAD37265	Human ABC1 transcr
5	1411.4	85.9	2910	24 AAD37267	Human ABC1 transcr
6	1146.4	69.8	1197	24 ABL58400	Human large ATP-bi
7	119.4	7.3	227	21 AAC09615	Human secreted pro
8	102.2	6.2	2910	24 AAD37267	Human ABC1 transcr
9	102.2	6.2	3231	24 AAD37265	Human ABC1 transcr

10	98	6.0	7260	22 AAD21326	Human ATP binding
11	98	6.0	7260	22 AAI70315	Human ATP binding
12	97	5.9	1643	22 AAF24681	Nucleotide sequenc
13	97	5.9	1643	22 AAF24703	Nucleotide sequenc
14	92.6	5.6	183999	22 AAF2831	Human ABC1 genomic
15	91	5.5	221	24 AAF3268	Human ABC1 gene ex
16	89.4	5.4	736	22 AAD37267	Human ABC1 transcr
17	89.4	5.4	1556	22 AAI8606	Human cDNA clone (
18	89	5.4	763	22 AAI8606	Human cDNA clone (
19	89	5.4	1750	22 AAI8606	Human cDNA clone (
20	77.2	4.7	7281	22 AAI8606	Human polynucleoti
21	75.4	4.6	7086	22 ABA09200	Human ABCA1 homolo
22	75.4	4.6	7086	22 AAK52667	Human polynucleoti
23	75	4.6	9854	22 AAS06121	Human ABC1 DNA seq
24	67	4.1	10442	22 AAF24680	Nucleotide sequenc
25	67	4.1	10442	22 AAF24702	Nucleotide sequenc
26	60	3.7	10474	22 AAF24685	Nucleotide sequenc
27	60	3.7	10474	22 AAF24685	Nucleotide sequenc
28	60	3.7	10474	22 AAF24707	Nucleotide sequenc
29	60	3.7	10474	22 AAF24708	Nucleotide sequenc
30	51.4	3.1	474	24 AAT73478	Bovine embryonic g
31	46.6	2.8	5116	24 AAT73478	Human chemically m
32	45	2.7	114955	20 AAT73478	Human chemically m
33	44.2	2.7	456	22 AAI83199	Human polynucleoti
34	43.8	2.7	385	22 AAI83199	Human polynucleoti
35	43.8	2.7	400	22 AAI83199	Human polynucleoti
36	43.4	2.6	401	22 AAI83199	Human polynucleoti
37	43.4	2.6	772	22 AAI83199	Human cDNA clone (
38	43.4	2.6	2339	22 AAI83199	Human cDNA clone (
39	43.4	2.6	10119	22 AAI83199	Human cDNA clone (
40	43.4	2.6	14983	22 AAI83199	Human immune/haema
41	43.4	2.6	14983	22 AAI83199	Human immune/haema
42	43.4	2.6	32190	22 AAI83199	Human immune/haema
43	43.4	2.6	32249	22 AAI83199	Human genomic DNA
44	43.4	2.6	39068	22 AAI83199	Human genomic DNA
45	43.4	2.6	39068	22 AAI83199	Human immune/haema

ALIGNMENTS

RESULT 1

AAF24681
ID AAF24681 standard; DNA; 1643 BP.

AC AAF24681:

XX
XX 20-APR-2001 (first entry)

DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

DE Human; adenosine triphosphate binding cassette protein 1; ABC1;

KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;

KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;

KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX WO200078972-A2.

XX 28-DEC-2000.

PD 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

ORIGINAL

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT useful for the development of agents for the treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Claim 1: Page 143-144; 215pp; English.
XX
CC The present sequence represents the 5' flanking region of the human
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABC1 is a pivotal
CC protein in the apolipoprotein-mediated mobilisation of intracellular
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
CC are useful for developing pharmaceutical agents for the treatment of
CC heart disease and other disorders associated with hypercholesterolemia
CC and atherosclerosis. The genes are useful for developing screening assays
CC to screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP: 370 A; 413 C; 457 G; 403 T: 0 other;
Query Match 100.0%; Score 1643; DB 22; Length 1643;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GAATTCCTTGTGCGGCTCCACATGCAGCTTCACAGGCGCTTGCTTGCTCTTATGGGTC 60
Db 1 GAATTCCTTGTGCGGCTCCACATGCAGCTTCACAGGCGCTTGCTTGCTCTTATGGGTC 60
QY 61 TGTCTGTAGTGTGTATAGAACACATGATGAGTGTGAGTACCTGGGCTTGAGCGTGGCGTGGACA 120
Db 61 TGTCTGTAGTGTGTATAGAACACATGATGAGTGTGAGTACCTGGGCTTGAGCGTGGCGTGGACA 120
QY 121 TCCTGTGTAGCTGTAGCATGGAGGGGGCTTGTCAGCTGAATGCTCTGTATCAGAGTGGTGGG 180
Db 121 TCCTGTGTAGCTGTAGCATGGAGGGGGCTTGTCAGCTGAATGCTCTGTATCAGAGTGGTGGG 180
QY 181 AGTTCTGGAATATGATGAGCTGGAGGTGGGAAGACAGTAGTACGCTTGGCGCACTCTCTC 240
Db 181 AGTTCTGGAATATGATGAGCTGGAGGTGGGAAGACAGTAGTACGCTTGGCGCACTCTCTC 240
QY 241 ATGCCACCTCATTTGCGCAAAACATCAGCTCAAACCTGTCAAGAGTCTAAATGTGAATCTG 300
Db 241 ATGCCACCTCATTTGCGCAAAACATCAGCTCAAACCTGTCAAGAGTCTAAATGTGAATCTG 300
QY 301 CCCTTCAAGTGGCTACAAAGATCTTTTGTCAAGGTAGGAGACCTTGHGGCGTCCACGT 360
Db 301 CCCTTCAAGTGGCTACAAAGATCTTTTGTCAAGGTAGGAGACCTTGHGGCGTCCACGT 360
QY 361 GCACCTCCAGGGCGCTGTTGGCCCTTCTACGGGTCTGCTCCAGTCTTCTATGAATCTC 420
Db 361 GCACCTCCAGGGCGCTGTTGGCCCTTCTACGGGTCTGCTCCAGTCTTCTATGAATCTC 420
QY 421 CCTTCAGGGCAGATTCATATTAGACTCTTACAGCTTTGACCTGAGTTTGGCCAGAATA 480
Db 421 CCTTCAGGGCAGATTCATATTAGACTCTTACAGCTTTGACCTGAGTTTGGCCAGAATA 480
QY 481 AGGTGCACATTTAGTTGTGCGTTGATCAATGACATTAATATTTACACATATGGTGTA 540
Db 481 AGGTGCACATTTAGTTGTGCGTTGATCAATGACATTAATATTTAGACATATGGTGTA 540
QY 541 GGCCTGCATTCCTACTCTTGCCCTTTTTTTTGGCCCTCCAGTGTTTTGGGTAAGTTTGGT 600
Db 541 GGCCTGCATTCCTACTCTTGCCCTTTTTTTTGGCCCTCCAGTGTTTTGGGTAAGTTTGGT 600
QY 601 CCCCCATAGCCCAAGGCAACAGATAGTTGGAGGTCTGGAGTGGCTGCTACATAATTATAC 660
Db 601 CCCCCATAGCCCAAGGCAACAGATAGTTGGAGGTCTGGAGTGGCTGCTACATAATTATAC 660


```
FT protein_bind /bound_moiety= "AP4"
FT 1498..1514 /tag= w
FT /bound_moiety= "LMO2-COM/MYOD/deltaEFL/E47"
FT 1524..1545 /tag= x
FT /bound_moiety= "ZID/deltaEFL"
FT 1597..1607 /tag= y
FT /bound_moiety= "DeltaEFL"
FT 1622..1627 /tag= z
FT /bound_moiety= "PPAR"
FT 1632..1637 /tag= aa
FT /bound_moiety= "PPAR"
FT 1685..1698 /tag= ab
FT /bound_moiety= "USE/NMYC/MYCKAX"
FT 1787..1797 /tag= ac
FT /bound_moiety= "DeltaEFL"
FT 1809..1819 /tag= ad
FT /bound_moiety= "DeltaEFL"
FT 1822..1833 /tag= ae
FT /bound_moiety= "SRY"
FT 1840..1850 /tag= af
FT /bound_moiety= "AP1"
FT 1942..1956 /tag= ag
FT /bound_moiety= "HNF3beta"
FT 1978..1985 /tag= ah
FT /bound_moiety= "NKX2.5"
FT 2008..2016 /tag= ai
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2019..2024 /tag= aj
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2051..2059 /tag= ak
FT /bound_moiety= "GATA"
FT 2104..2111 /tag= al
FT /bound_moiety= "SOX5"
FT 2114..2152 /tag= am
FT /bound_moiety= "SRY/HNF3beta"
FT 2221..2228 /tag= an
FT /bound_moiety= "MZFL"
FT 2234..2249 /tag= ao
FT /bound_moiety= "IK2/NFkappaB/CREL"
FT 2259..2272 /tag= ap
FT /bound_moiety= "LMO2COM/GATA"
FT 2289..2306 /tag= aq
FT /bound_moiety= "MZFL/SRY"
FT 2313..2318 /tag= ar
FT /bound_moiety= "PPAR"
FT 2321..2326 /tag= as
FT /bound_moiety= "PPAR"
FT 2335..2342 /tag= at
FT /bound_moiety= "MZFL"
```

```
FT protein_bind 2361..2384 /tag= au
FT /bound_moiety= "HNF3beta/SRY/EV1"
FT 2426..2433 /tag= av
FT /bound_moiety= "MZFL"
FT 2442..2451 /tag= aw
FT /bound_moiety= "AP4"
FT 2455..2466 /tag= ax
FT /bound_moiety= "SRY"
FT 2491..2498 /tag= ay
FT /bound_moiety= "STAT"
FT 2524..2534 /tag= az
FT /bound_moiety= "STAT/PPAR"
FT 2536..254 /tag= ba
FT /bound_moiety= "PPAR"
FT 2589..2600 /tag= bb
FT /bound_moiety= "AP2"
FT 2610..2617 /tag= bc
FT /bound_moiety= "MZFL"
FT 2634..2648 /tag= bd
FT /bound_moiety= "LMO2COM/MYOD/E47"
FT 2657..2672 /tag= be
FT /bound_moiety= "RRE31"
FT 2680..2698 /tag= bf
FT /bound_moiety= "MZFL/CMYB"
FT 2728..2740 /tag= bg
FT /bound_moiety= "SP1/GC"
FT 2743..2757 /tag= bh
FT /bound_moiety= "USE/NMYC/ARNT"
FT 2758..2773 /tag= bi
FT /bound_moiety= "NFE2A1"
FT 2774..2787 /tag= bj
FT /bound_moiety= "XFDL/HFH"
FT 2794..2806 /tag= bk
FT /bound_moiety= "GC/SP1/MZF1"
```

Query Match 85.9%; Score 1411.4; DB 24; Length 2910;

Best Local Similarity 98.0%; Pred. NO. 0;

Matches 1541; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

QY 16 GCCTCCAGATGCACCTCCAGGGCCCTGCTGG--CCTCTCATGGCTCTCTCTCTGAGTGT 73

DB 1345 GCCTCCAGATGCACCTCCAGGGCCCTGCTGGCCCTCTCTATGCTCTCTCTGAGTGT 1404

QY 74 GATAGAACACCTGATGTGATGACCTGGGCTTGAG--CGTGGCCTGGAGATCCCTGTGACTG 132

DB 1405 GATAGAACACCTGATGTGATGACCTGGGCTTGAGCCCTGGAGATCCCTGTGACTG 1464

QY 133 TAGCATGGAGGGGGTGTGT--CAGCTCAATGTCTGTATGAGGTGGTGGAGTCTGGAGT 191

DB 1465 TAGCATGGAGGGGGTGTGTGTGTCAGCTGAATGTCTGATGATGAGGTGGTGGAGTCTGGAGT 1524

QY 192 ATGATGGAGCTGGAGGTGGAGAGAGTAGGCTGGGGCAGCTCTCTCATGCCACCTCA 251

DB 1525 ATGATGGAGCTGGAGGTGGAGAGAGTAGGCTGGGGCAGCTCTCTCATGCCACCTCA 1584

QY 252 TTCTGGCCAAACACAGGTCAAACCTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAAGT 311

Query Match 59.8%; Score 1146.4; DB 24; Length 1197;
Best Local Similarity 99.7%; Pred. No. 1.2e-306;
Matches 1180; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 460 ACCTGAGTTTGGCCAGATAAAGGTACATTTAGTTTGGTTGATGATGACATTA 519
DB 1 ACCTGAGTTTGGCCAGATAAAGGTACATTTAGTTTGGTTGATGATGACATTA 60
QY 520 TATTTAGACATATGGTGTGAGCCCTGCATCTACTCTTTCCTTTTTCCTTCC 579
DB 61 TATTTAGACATATGGTGTGAGCCCTGCATCTACTCTTTCCTTTTTCCTTCC 120
QY 580 AGTGTCTTGGGTAGTTTGTGTCCTCCCTACAGCCAAAGGCAAGATGAGTGT 639
DB 121 AGTGTCTTGGGTAGTTTGTGTCCTCCCTACAGCCAAAGGCAAGATGAGTGT 179
QY 640 GGAGTGGCTACATAATTTACAGACATGCAATTTCTGCTGGCTGCACATTC 699
DB 180 GGAGTGGCTACATAATTTACAGACATGCAATTTCTGCTGGCTGCACATTC 239
QY 700 AACTTAATACAACTCTCTGTTTATACAGGAGGCTGATCAATATGAATTA 759
DB 240 AACTTAATACAACTCTCTGTTTATACAGGAGGCTGATCAATATGAATTA 299
QY 760 AAGGGGCTGGTCCCATATGTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 819
DB 300 AAGGGGCTGGTCCCATATGTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 358
QY 820 TGGCTCTCTCTCTCAATTTATGAAGAGAGCAGTAGATGTTCTCTCGGCTCT 879
DB 359 TGGCTCTCTCTCTCAATTTATGAAGAGAGCAGTAGATGTTCTCTCGGCTCT 418
QY 880 AGGACCTGGGAGCTCAGGCTGGGAATCTCAAGAGCAGTAGTGGCTATCAAA 939
DB 419 AGGACCTGGGAGCTCAGGCTGGGAATCTCAAGAGCAGTAGTGGCTATCAAA 478
QY 940 AAGTCCAGGTTTGGGGGGGAAACAAAGCAGCCCATACCCAGAGGACTTCCG 999
DB 479 AAGTCCAGGTTTGGGGGGGAAACAAAGCAGCCCATACCCAGAGGACTTCCG 538
QY 1000 CCCCTCACCCAGCTAGGCTTTGAAAGGAAACAAAGACAGCAAAATGATGGGT 1059
DB 539 CCCCTCACCCAGCTAGGCTTTGAAAGGAAACAAAGACAGCAAAATGATGGGT 598
QY 1060 CCGTGGAGGATTCAGCCTAGAGCTCTCTCTCCCAATTCCTCTCGGCTGAGAAA 1119
DB 599 CCGTGGAGGATTCAGCCTAGAGCTCTCTCTCCCAATTCCTCTCGGCTGAGAAA 557
QY 1120 CTAAACAGGAAACAAATTCGGNAACAGGATTTAGAGGAGCAAAATTCACCTGGTG 1179
DB 658 CTAAACAGGAAACAAATTCGGNAACAGGATTTAGAGGAGCAAAATTCACCTGGTG 717
QY 1180 CCGTGGCTGGCGGAACCTGAGCTAGAGTCTGGGGCGAGCCCGAGCCGAGCGCTT 1239
DB 718 CCGTGGCTGGCGGAACCTGAGCTAGAGTCTGGGGCGAGCCCGAGCCGAGCGCTT 777
QY 1240 CCGCGCGCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 1299
DB 778 CCGCGCGCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 837
QY 1300 AGACACTGCTGATGCTCCAGCCCAACCCACCCACCTCCCGCCACCTCCCTAGATGTG 1359
DB 838 AGACACTGCTGATGCTCCAGCCCAACCCACCCACCTCCCGCCACCTCCCTAGATGTG 897
QY 1360 TCGTGGCGGCTGAACCTGCGCGCTTTAAGGGCGGGCGCGCGCGCGCGCTTTCG 1419
DB 898 TCGTGGCGGCTGAACCTGCGCGCTTTAAGGGCGGGCGCGCGCGCGCTTTCG 957
QY 1420 TGAGTGAATGAACTACATAACAGAGCGCGGGAACGCGCGCGCGGAGGAGGAGCACAG 1479
DB 958 TGAGTGAATGAACTACATAACAGAGCGCGGGAACGCGCGCGCGGAGGAGGAGCACAG 1017

QY 1480 GCTTTGACCCAGTAGTAACCTCTGCGCTGCGAGCCGAATCTATAAAGGAAGTACTGTC 1539
DB 1018 GCTTTGACCCAGTAGTAACCTCTGCGCTGCGAGCCGAATCTATAAAGGAAGTACTGTC 1077
QY 1540 CGSCAAACCCCGCTAATTCGCGAGCGAGTGAAGTGGGCGCGGAGCCGCGAGAGCGAGC 1599
DB 1078 CGSCAAACCCCGCTAATTCGCGAGCGAGTGAAGTGGGCGCGGAGCCGCGAGAGCGAGC 1137
QY 1600 CGACCTTCTCTCCCGGCTGCGGCGAGCGAGCGCGGCGAGCGCTC 1643
DB 1138 CGACCTTCTCTCCCGGCTGCGGCGAGCGAGCGCGGCGAGCGCTC 1181

RESULT 7
AAC09615
ID AAC09615 standard; cDNA; 227 bp.
XX AAC09615:
XX AC
XX AC
DT 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 13690.
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
PD 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
PR (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI: 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 13690; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX Sequence 227 bp; 44 A; 65 C; 73 G; 45 T; 0 other;

Query Match 7.3%; Score 119.4; DB 21; Length 227;
Best Local Similarity 99.2%; Pred. No. 7.1e-23;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1523 ATAAAGGAAGTACTAGTCCCGGCAAAACCCGCTAATTCGAGCGAGAGTGAAGTGGGCGCG 1582
DB 1 ATAAAGGAAGTACTAGTCCCGGCAAAACCCGCTAATTCGAGCGAGAGTGAAGTGGGCGCG 60

```
Oy 1593 GACCCGAGAGCCGAGCCGACCTCTCTCCGGCTCCGGCAGGGCAGGGCGGAGCT 1642
      |||||
Db 61 GACCCGAGAGCCGAGCCGACCTCTCTCCGGCTCCGGCAGGGCAGGGCGGAGCT 120
Oy 1643 C 1643
Db 121 C 121

RESULT 8
AAD37267/C
ID AAD37267 standard; DNA; 2910 bp.
XX
AC AAD37267;
DT 21-AUG-2002 (first entry);
DE Human ABC1 transcription regulatory DNA #3.
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX Homo sapiens.
OS
FH Key
FH Location/Qualifiers
FT 12...23
FT /tag= a
FT /bound_moiety= "LMO2COM/MYOD"
FT protein_bind
FT 97...107
FT /tag= b
FT /bound_moiety= "DeltaEF1"
FT protein_bind
FT 110...125
FT /tag= c
FT /bound_moiety= "S3/NKX2.5"
FT protein_bind
FT 196...211
FT /tag= d
FT /bound_moiety= "S8"
FT protein_bind
FT 228...237
FT /tag= e
FT /bound_moiety= "GATA"
FT protein_bind
FT 399...410
FT /tag= f
FT /bound_moiety= "IK2"
FT protein_bind
FT 412...420
FT /tag= g
FT /bound_moiety= "LYF1"
FT protein_bind
FT 528...539
FT /tag= h
FT /bound_moiety= "LMO2COM/MYOD/DeltaEF1"
FT protein_bind
FT 549...556
FT /tag= i
FT /bound_moiety= "LYF1"
FT protein_bind
FT 558...568
FT /tag= j
FT /bound_moiety= "DeltaEF1"
FT protein_bind
FT 590...596
FT /tag= k
FT /bound_moiety= "LXX2.5"
FT protein_bind
FT 608...620
FT /tag= l
FT /bound_moiety= "NFY/CAAT"
FT protein_bind
FT 708...715
FT /tag= m
FT /bound_moiety= "MZFl"
FT protein_bind
FT 723...730
FT /tag= n
FT /bound_moiety= "MZFl"
FT protein_bind
FT 771...785
FT /tag= o
FT /bound_moiety= "HNF3beta"
FT protein_bind
FT 803...812
FT /tag= p
FT /bound_moiety= "IK2/NFkappaB/CREL"
FT protein_bind
FT 831...837
FT /tag= r
FT /bound_moiety= "NKX2.5"
FT protein_bind
FT 1075...1089
FT /tag= s
FT /bound_moiety= "GATA"
FT protein_bind
FT 1173...1188
FT /tag= t
FT /bound_moiety= "LXRalpha/deltaEF1"
FT protein_bind
FT 1189...1209
FT /tag= u
FT /bound_moiety= "DeltaEF1/LYF1/IK2"
FT protein_bind
FT 1483...1491
FT /tag= v
FT /bound_moiety= "A24"
FT protein_bind
FT 1498...1514
FT /tag= w
FT /bound_moiety= "LMO2-COM/MYOD/deltaEF1/B47"
FT protein_bind
FT 1524...1545
FT /tag= x
FT /bound_moiety= "ZID/deltaEF1"
FT protein_bind
FT 1597...1607
FT /tag= y
FT /bound_moiety= "DeltaEF1"
FT protein_bind
FT 1622...1627
FT /tag= z
FT /bound_moiety= "PPAR"
FT protein_bind
FT 1632...1637
FT /tag= aa
FT /bound_moiety= "PPAR"
FT protein_bind
FT 1685...1698
FT /tag= ab
FT /bound_moiety= "USF/NMYC/MYCMAX"
FT protein_bind
FT 1787...1797
FT /tag= ac
FT /bound_moiety= "DeltaEF1"
FT protein_bind
FT 1809...1819
FT /tag= ad
FT /bound_moiety= "DeltaEF1"
FT protein_bind
FT 1822...1833
FT /tag= ae
FT /bound_moiety= "SRY"
FT protein_bind
FT 1840...1850
FT /tag= af
FT /bound_moiety= "API"
FT protein_bind
FT 1842...1856
FT /tag= ag
FT /bound_moiety= "HNF3beta"
FT protein_bind
FT 1978...1985
FT /tag= ah
FT /bound_moiety= "NKX2.5"
FT protein_bind
FT 2008...2016
FT /tag= ai
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT protein_bind
FT 2019...2024
FT /tag= aj
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT protein_bind
FT 2051...2059
FT /tag= ak
FT /bound_moiety= "GATA"
FT protein_bind
FT 2104...2111
FT /tag= al
FT /bound_moiety= "SOX5"
FT protein_bind
FT 2114...2152
FT /tag= am
FT /bound_moiety= "SYR/HFH/HNF3beta"
FT protein_bind
FT 2221...2228
FT /tag= an
FT /bound_moiety= "MZFl"
FT protein_bind
FT 2234...2249
FT /tag= ao
FT /bound_moiety= "IK2/NFkappaB/CREL"
FT protein_bind
FT 2259...2272
```

RESULT 9	
AA037265/C	
ID	AA037265 standard; DNA; 3231 BP.
XX	
XX	AA037265;
XX	
XX	AC
XX	
XX	21-AUG-2002 (first entry)
XX	
XX	Human ABC1 transcription regulatory DNA #1.
DE	
XX	Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
XX	KW
XX	cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX	
XX	Homo sapiens.
OS	
XX	W0200183746-A2.
PN	
XX	
XX	08-NOV-2001.
PD	
XX	
XX	02-MAY-2001; 2001WO-EP05488.
2F	
XX	
XX	02-MAY-2000; 2000US-201250P.
PR	
XX	
XX	(AVET) AVENTIS PHARMA SA.
XX	
XX	Resier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe F;
XX	Brewer B, Duverger N, Remaley A, Santamarina-Pojo S;
PI	
XX	
XX	WPI: 2002-154404/20.
DR	
XX	
XX	Isolated nucleic acid useful for modifying the ATP-binding cassette 1
PT	(ABC1) and screening for candidate modulatory compounds or substances
PT	.
XX	
XX	Claim 1; Page 130-131; 152pp; English.
PS	
XX	
XX	The invention relates a nucleic acid which is capable of regulating the
CC	transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
CC	casual gene for pathologies linked to a dysfunctioning of cholesterol
CC	metabolism, including diseases such as atherosclerosis. Polynucleotid
CC	of the invention are used to screen candidate molecules or substances
CC	that are capable of modulating the transcription of the ABC1 gene. The
CC	are used in antisense therapy. Compositions comprising sequences of t
CC	invention are used to treat hypercholesterolaemia and atherosclerosis
CC	The present sequence is human ABC1 transcription regulating DNA.
XX	
XX	Sequence 3231 BP; 809 A; 773 C; 876 G; 773 T; 0 other;
SO	

Query Match 6.2%; Score 102.2; DB 24; Length 3231;
Best Local Similarity 55.9%; Pread No. 1.6e-17;
Matches 164; Conservative 0; Mismatches 93; Indels 2; Gaps 1;

The invention relates a nucleic acid which is capable of regulating the transcription of human ATP-binding cassette 1 (ABCL1) gene, which is a causal gene for pathologies linked to a dysfunctioning of cholesterol metabolism, including diseases such as atherosclerosis. Polynucleotides of the invention are used to screen candidate molecules or substances that are capable of modulating the transcription of the ABCL1 gene. They are used in antisense therapy. Compositions comprising sequences of the invention are used to treat hypercholesterolemia and atherosclerosis. Inverted sequences is human ABCL1 transcription regulating DNA.

301 CCCTCAAGTGGCTACAAAGGTAICTTTTC-CAAGGTAGGAGACCTTTGGCCCTCCACGT 360
 1760 CCCTGAAGGATTCATAGAAGACICAGGACAGACCCCGTAGAAGGCGCCAAAGCGCCCTG 1701
 361 GCATCTTCAGGCGCTGCTTGCCCTCTCTCTAGGGGTCTGCTCTGAGTCTCTATGAATCTC 420
 1700 GAAGTGACGCTGGAGGC--CACAAAGGCTCTCTACCTTGACAAAGATACCTTTGTAGCCAC 1643
 421 CCTTCAGGCGAGATTCATPATTAGACTCTTCACAGCTTTGACTGAGTTTGGCCAGATA 480
 1642 CTTGAAGGCGAGATTCACATTTAGACTCTTCACAGCTTTGACTGAGTTTGGCCAGATA 1583
 481 AGGTGACAT 489
 1582 AGGTGACAT 1574

RESULT 10
 AAD21326
 ID AAD21326 standard; DNA; 7260 BP.
 XX
 AC AAD21326;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human ATP binding cassette transporter 1 (ABCI) gene.
 XX
 KW Human: ATP binding cassette transporter 1; ABC1; coronary heart disease;
 KW dermatological; atherosclerosis; cardiovascular; inflammatory disease;
 KW psoriasis; lipid disorder; antibacterial; septic shock; gene therapy;
 KW immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 321..7106
 FT /*tag= a
 FT /product= "Human ABC1 protein"
 FT
 FT
 PN EP1136552-A1.
 XX
 PD 26-SEP-2001.
 XX
 PF 20-MAR-2000; 2000EP-0105820.
 XX
 PR 20-MAR-2000; 2000EP-0105820.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Schmitz G, Bodzioch M;
 XX
 DR WPI; 2001-640389/74.
 DR P-PSDB; AAE13022.
 XX
 PT New adenosine triphosphate binding cassette transporter-1 gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases
 XX
 PS Example 1; Fig 1; 48pp; English.
 XX
 CC The invention relates to four common polymorphisms in the gene encoding
 CC ATP-binding cassette transporter-1 (ABCI). ABC1 is associated with
 CC decreased ApoA-I mediated efflux of cholesterol. The polymorphisms in
 CC ABC1 directly affects cellular lipid homeostasis, which is a key factor
 CC in the atherogenic processes. The ABC1 polymorphisms are useful for
 CC diagnosing and treating lipid disorders, cardiovascular diseases
 CC (coronary heart disease, atherosclerosis) and inflammatory diseases
 CC (psoriasis, lupus erythematosus). The identification of ABC1 as a
 CC transporter for interleukin-beta (IL-beta) identifies this gene as
 CC a candidate for treatment of inflammatory diseases including rheumatoid
 CC arthritis and septic shock. The present sequence is human ABC1 gene.

SQ Sequence 7260 BP; 1834 A; 1755 C; 1905 G; 1756 T; 0 other;
 Query Match 6.0%; Score 98; DB 22; Length 7260;
 Best Local Similarity 100.0%; Pred. No. 3.4e-16;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1546 AAACCCCGCTAATTCGAGCGAGAGTGTGGGCGCGGACCCGAGAGCCGAGCCGACCC 1605
 DB 1 AAACCCCGCTAATTCGAGCGAGAGTGTGGGCGCGGACCCGAGAGCCGAGCCGACCC 60
 QY 1606 TTCTCTCCCGGCTCCGCGACGCGGCGGCGGAGCTC 1643
 DB 61 TTCTCTCCCGGCTCCGCGACGCGGCGGCGGAGCTC 98

RESULT 11
 AAI70315
 ID AAI70315 standard; cDNA; 7260 BP.
 XX
 AC AAI70315;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human ATP binding cassette transporter 1 (ABCI) cDNA.
 XX
 KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;
 KW cholesterol; cardiovascular disease; inflammatory disease;
 KW antiinflammatory; antilipemic; antipsoriatic; dermatological;
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
 KW polymorphism; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 321..7106
 FT /*tag= a
 FT CDS 501..7106
 FT /*tag= b
 FT /note= "alternative open reading frame of AAI70314"
 FT variation replace(975,A)
 FT /*tag= c
 FT variation replace(1515,C)
 FT /*tag= d
 FT variation replace(2969,G)
 FT /*tag= e
 FT variation replace(3836,C)
 FT /*tag= f
 XX
 PN EP1136554-A1.
 XX
 PD 26-SEP-2001.
 XX
 PF 24-MAR-2000; 2000EP-0105401.
 XX
 PR 24-MAR-2000; 2000EP-0106401.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Schmitz G, Bodzioch M;
 XX
 DR WPI; 2001-640389/74.
 DR P-PSDB; AAM50228.
 XX
 PT New adenosine triphosphate binding cassette transporter gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases
 XX
 PS Disclosure; Page 26-28; 41pp; English.
 XX
 CC The present sequence is that of cDNA encoding the human adenosine
 CC triphosphate (ATP) binding cassette transporter 1 (ABCI) protein
 CC (see AAM50227). The sequence includes an extended open reading
 CC frame (ORF) to that provided by the sequence in AAI70314, using

an alternative ATG codon as initiation codon and thereby adding an extra 40 N-terminal amino acids to the encoded ABC1 protein (see CC AAM50228). The invention provides 4 common polymorphisms in the CC ABC1 gene. These were identified by sequencing the ABC1 gene in CC different Tangier kindreds. In the variant genes (numbered as in CC AAI70314), G is changed to A at position 596. T is changed to C at CC position 1136. A is changed to G at position 2599 or G is changed CC to C at position 3456, or any combination of these. All of these CC polymorphisms alter the amino acid sequence of ABC1 and therefore CC may affect its function. The 2 most common polymorphisms (G596A) CC and A2599G) are both associated with a decreased in vitro ApoA-I CC mediated efflux of cholesterol from mononuclear phagocytes, a CC feature typical of Tangier disease. 3 Of the variants (G596A, CC A2599G and G3456C) are significantly increased in a population of CC men having low high density lipoprotein-cholesterol levels and CC established coronary heart disease (CHD) relative to CHD-free CC control subjects. The use of the provided ABC1 polymorphisms for CC the diagnosis and treatment of lipid disorders, cardiovascular CC diseases, and inflammatory diseases (e.g. psoriasis, lupus CC erythematoses) is claimed. Modulation of ABC1 transcripts or CC proteins by antisense or ribozyme technology or RNA decoys is also CC claimed.

XX Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 6.0%; Score 98; DB 22; Length 7260;

Best Local Similarity 100.0%; Pred. No. 3.4e-16;

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1546 AAACCCCGTAATTCGAGCGAGAGTGTAGTGGCGGCGGACCCGAGCGGAGCCGACCC 1605

DB 1 AAACCCCGTAATTCGAGCGAGAGTGTAGTGGCGGCGGACCCGAGCGGAGCCGACCC 50

QY 1606 TTCTCTCCCGGCTGCGCGAGCGAGCGGCGGAGCTC 1643

DB 61 TTCTCTCCCGGCTGCGCGAGCGAGCGGCGGAGCTC 98

RESULT 12

AAF24681/c

ID AAF24681 standard; DNA; 1643 BP.

XX AAF24681;

XX 20-APR-2001 (first entry)

Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;

KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;

KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;

KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-); CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/4.

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide, useful for the development of agents for the treatment of heart disease

PT and other disorders associated with hypercholesterolemia and atherosclerosis

XX Claim 1; Page 143-144; 215pp; English.

XX The present sequence represents the 5' flanking region of the human adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1 resides in cell membranes and utilizes ATP hydrolysis to transport a wide variety of substrates across the plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated mobilisation of intracellular cholesterol stores. ABC1 is defective in Tangier disease, a genetic disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins are useful for developing pharmaceutical agents for the treatment of heart disease and other disorders associated with hypercholesterolemia and atherosclerosis. The genes are useful for developing screening assays to screen for compounds that regulate the expression of genes associated with cholesterol transport. The genes and proteins are also useful for are also useful as diagnostic indicators of cardiovascular disease and other disorders associated with hypercholesterolemia.

XX Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 5.9%; Score 97; DB 22; Length 1643;

Best Local Similarity 61.8%; Pred. No. 3e-16;

Matches 154; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 241 ATGCCACCTCAATTCGCCCAAACTCAGGTCAAACTGGAAGAGTCTAAATGTGAATCTG 300

DB 489 ATGTCACCTTAATTCGCCCAAACTCAGGTCAAACTGGAAGAGTCTAAATGTGAATCTG 430

QY 301 CCCTTCAAGGTGGGTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCCTCCAGGT 360

DB 429 CCCTCAAGGAGATTCATAGAGACTCAGGACAGACCCGTAGAGAGCCCAAGCGGCC 370

QY 361 GCACCTTCAGGGCGGTGTTGGCCCTCTTCTACGGGTCTCTCTGAGTCTTCTATGATCTC 420

DB 369 TGAAGATGCACGTGGAGGCCCAAGGTCTCTTACCTTGACAAAGATACCTTTGTAGCCAC 310

QY 421 CCCTCAGGCGGAGATTCATATTAGACTCTTCACAGTTTGACCTGATTTGGCCAGATA 480

DB 309 CTTGAAGGCGAGATTCACATTAGACTCTTCACAGTTTGACCTGATTTGGCCAGATG 250

QY 481 AGGTGACAT 489

DB 249 AGGTGGCAT 241

RESULT 13

AAF24703/c

ID AAF24703 standard; DNA; 1643 BP.

XX AAF24703;

XX 20-APR-2001 (first entry)

Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;

KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;

KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;

XX atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078971-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16591.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.
XX (CVTH-) CV THERAPEUTICS INC.
PA (UNIV) UNIV WASHINGTON.
XX
XX LAW: RM, Wade D, Oram JF, Garvin M;
XX WP: 2001-137811/14.
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT polynucleotides and polypeptides, useful for treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
XX Disclosure; Page 138-139; 21pp; English.
XX
XX The present sequence represents the 5' flanking region of the human
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC resides in cell membranes and utilizes ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABC1 is a pivotal
CC protein in the apolipoprotein-mediated mobilization of intracellular
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC disorder characterized by abnormal HDL-cholesterol metabolism. The ABC1
CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
CC are useful for developing pharmaceutical agents for the treatment of
CC heart disease and other disorders associated with hypercholesterolemia
CC and atherosclerosis. The genes are useful for developing screening assays
CC to screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC with cholesterol transport. The genes and proteins are also useful for
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 5.9%; Score 97; DB 22; Length 1643;
Best Local Similarity 61.8%; Pred. No. 3e-16; Mismatches 0; Gaps 0;
Matches 154; Conservative 0; Indels 95;

QY 241 ATGCGACCTCATCTCGCCAAACTCAGTCAAACTGCAAGAGTCTAAATGTGAATCTG 300
DB 489 ATGTCACCTTATCTCGCCAAACTCAGTCAAACTGCAAGAGTCTAAATGTGAATCTG 430
QY 301 CCTTCAAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGACCTTGTGGCCCTCCACGT 360
DB 429 CCTGAAGGAGAGATTATAGAGACTCAGGACAGACCCCTAGAGAGGCCAAGCAGGCC 370
QY 361 GCACCTCCAGGGCCCTGCTTGGCTCTTCTACGGGTCTGCTTCTATGATCTC 420
DB 369 TGAAGTGCACGTGGAGGCCACAAAGGCTCTCTACCTTGACAAAGATACCTTGTAGCCAC 310
QY 421 CCTTACGGGCAGATTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAA7A 480
DB 309 CTTGAAGGCGAGATTACATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAA7G 250
QY 481 AGGTGACAT 489
DB 249 AGGTGGCAT 241

RESULT 14
AAF92831/c
ID AAF92831 standard; DNA; 183999 BP.
XX
XX AAF92831;
XX
XX 17-MAY-2001 (first entry)
XX Human ABC1 genomic DNA.
XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX Homo sapiens.
XX

PN W0200115676-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-IB01492.
XX
XX 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX (UYER-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WP: 2001-244356/25.
XX
XX Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
XX Claim 8; Fig 1; 317pp; English.
XX
XX The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;
SQ
Query Match 5.6%; Score 92.6; DB 22; Length 183999;
Best Local Similarity 63.1%; Pred. No. 5.3e-14;
Matches 159; Conservative 0; Mismatches 90; Indels 3; Gaps 1;
QY 241 ATGCGACCTCATCTCGCCAAACTCAGTCAAACTGCAAGAGTCTAAATGTGAATCTG 300
DB 27714 ATGTCACCTTATCTCGCCAAACTCAGTCAAACTGCAAGAGTCTAAATGTGAATCTG 27655
QY 301 CCTTCAAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGACCTTGTGGCCCTCCACGT 360
DB 27654 CCTGAAGGAGAGATTATAGAGACTCAGGACAGACCCCTAGAGAGGCCAAGCAGG 27595
QY 361 GCACCTCCAGGGCCCTGCTTGGCTCTTCTACGGGTCTGCTTCTATGATGAAT 417
DB 27594 CCTGGAAGTGCACGTGGAGGCCACAAAGGCTCTCTACCTTGACAAAGATACCTTGTAGC 477
QY 418 CTCCCTTACGGGCAGATTTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGA 477
DB 27534 CACCTTGAAGGCGAGATTTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGA 27475
QY 478 ATAGGTGACAT 489
DB 27474 ATAGGTGGCAT 27463

RESULT 15
AAD37268
ID AAD37268 standard; DNA; 221 BP.
XX
XX AAD37268;
XX
XX 21-AUG-2002 (first entry)
XX Human ABC1 gene exon 1A.
XX Human: ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
XX

KW cholesterol metabolism; hypercholesterolaemia; antisense therapy;
KW exon 1A; ds.
XX Homo sapiens.

XX WO200183746-A2.

XX PD 08-NOV-2001.

XX PF 02-MAY-2001; 23C1WO-EP05488.

XX PR 02-MAY-2000; 2000US-201280P.

XX PA (AVET) AVENTIS PHARMA SA.

XX PI Rosier-Ventus M, Prades C, Lemoine C, Naudin L, Deneffe P;

XX PI Brewer B, Duverger N, Remaley A, Santamarina-Pojo S;

XX DR WPI; 2002-154404/20.

XX PT Isolated nucleic acid useful for modifying the ATP-binding cassette 1
PT (ABCL) and screening for candidate modulatory compounds or substances

XX PS Claim 4; Page 132; 152pp; English.

XX CC The invention relates a nucleic acid which is capable of regulating the
CC transcription of human ATP-binding cassette 1 (ABCL) gene, which is a
CC casual gene for pathologies linked to a dysfunctioning of cholesterol
CC metabolism, including diseases such as atherosclerosis. Polynucleotides
CC of the invention are used to screen candidate molecules or substances
CC that are capable of modulating the transcription of the ABCL gene. They
CC are used in antisense therapy. Compositions comprising sequences of the
CC invention are used to treat hypercholesterolaemia and atherosclerosis.
XX CC The present sequence is human ABCL gene exon 1A.

XX SQ Sequence 221 BP; 44 A; 62 C; 73 G; 42 T; 0 other;

Query Match 5.5%; Score 91; DB 24; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GTAATTGGCAGCGAGAGTGAATGGGGCGGGAGCCCGCAGAGCCGCCCTTCTCTC 1612

Db 1 GTAATTGGCAGCGAGAGTGAATGGGGCGGGAGCCCGCAGAGCCGCCCTTCTCTC 60

QY 1613 CCGGGCTCGGCAGGGCAGGGCGGGGAGGCTC 1643

Db 61 CCGGGCTCGGCAGGGCAGGGCGGGGAGGCTC 91

Search completed: February 3, 2003, 16:20:03
Job time : 972.868 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:56:40 ; Search time 16,5211 Seconds
(without alignments)
6534.056 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1292_1643

Perfect score: 352

Sequence: 1 ggaacctgaacacctgctg.....cagggcagggcgaggagctc 352

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	35	9.9	7218	1	US-08-232-463-14
C 2	34.2	9.7	1610	4	US-09-276-331-56
C 3	32.6	9.3	3250	4	US-06-122-126B-14
C 4	32.2	9.1	16389	4	US-09-741-154-3
C 5	31.6	9.0	1835	4	US-09-485-549-1
C 6	31.4	8.9	1817	2	US-08-870-518-8
C 7	31.4	8.9	2060	3	US-09-205-143-1
C 8	31.4	8.9	81001	4	US-09-750-580-1
C 9	31.2	8.9	975	1	US-08-671-525B-9
C 10	31.2	8.9	975	1	US-08-672-109B-9
C 11	31.2	8.9	975	1	US-08-842-045-9
C 12	31.2	8.9	975	1	US-08-842-238-9
C 13	31.2	8.9	975	3	US-08-629-335B-9
C 14	31.2	8.9	978	3	US-08-706-281A-17
C 15	31.2	8.9	978	4	US-09-097-231-17
C 16	31.2	8.9	441529	4	US-09-103-840A-1
C 17	30.6	8.7	1969	4	US-09-038-628-1
C 18	30.6	8.7	17425	4	US-09-511-625B-5
C 19	30.6	8.7	43280	2	US-08-804-227C-1
C 20	30.2	8.6	1154	3	US-08-651-136C-7
C 21	30.2	8.6	1154	4	US-09-228-911A-7
C 22	30	8.5	966	3	US-08-469-318-68
C 23	30	8.5	966	3	US-08-469-318-70
C 24	30	8.5	966	3	US-08-469-318-71
C 25	30	8.5	966	3	US-08-468-609A-68
C 26	30	8.5	966	3	US-08-468-609A-70
C 27	30	8.5	966	3	US-08-468-609A-71

C 28	30	8.5	966	4	US-08-446-872A-68	Sequence 68, Appl
C 29	30	8.5	966	4	US-08-446-872A-70	Sequence 70, Appl
C 30	30	8.5	966	4	US-08-446-872A-71	Sequence 71, Appl
C 31	30	8.5	966	4	US-08-762-227A-68	Sequence 68, Appl
C 32	30	8.5	966	4	US-08-762-227A-70	Sequence 70, Appl
C 33	30	8.5	966	4	US-08-762-227A-71	Sequence 71, Appl
C 34	30	8.5	966	5	PCT-US95-01185-68	Sequence 68, Appl
C 35	30	8.5	966	5	PCT-US95-01185-70	Sequence 70, Appl
C 36	30	8.5	966	5	PCT-US95-01185-71	Sequence 71, Appl
C 37	30	8.5	35100	1	US-08-306-831B-19	Sequence 19, Appl
C 38	30	8.5	35100	5	PCT-US93-06251-19	Sequence 19, Appl
C 39	29.8	8.5	8310	3	US-08-870-126-11	Sequence 11, Appl
C 40	29.8	8.5	8310	4	US-09-445-247-11	Sequence 11, Appl
C 41	29.8	8.5	14985	1	US-08-652-972A-6	Sequence 6, Appl
C 42	29.8	8.5	14985	5	PCI-US86-06231A-6	Sequence 6, Appl
C 43	29.6	8.4	2224	3	US-09-109-204-6	Sequence 6, Appl
C 44	29.6	8.4	68750	3	US-09-335-409-1	Sequence 1, Appl
C 45	29.6	8.4	68750	4	US-09-568-102-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; City: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BERT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)683-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: p72gpt-F1s
US-08-232-463-14

Query Match 9.9% ; Score 35; DB 1; Length 7218;


```
Best Local Similarity 8.0%; Pred. No. 0.76;
Matches 17; Conservative 113; Mismatches 83; Indels 0; Gaps 0;

QY 130 GAGTACGACGACTACATAACAGACGGCGGAGACGGCGGGGAGGAGGAGACAGG 189
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1245 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1186
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 CTITGACGAGTACCTCGCTCGGTGCGGCGGAGTCTATATAAGAACTAGTCCC 249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1185 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1126
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 GCGAAACACCCGTAATGCGAGCGAGAGTACGCGGGCGGAGCCGCGAGACGCGC 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1125 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1066
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 GACCCCTCTCTCCCGGCTCGCGGCGGAGG 342
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1065 TCGAAGCTCCCTCGACCTGCACCAAGCTCGG 1033

RESULT 2
US-09-276-531-56
; Sequence 56, Application US/09276531
; Patent No. 5183988
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Koopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1610 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT08
; CLONE: -395833
US-09-276-531-56
```

```
Query Match: 9.7%; Score 34.2; DB 4; Length 1610;
Best Local Similarity 54.5%; Pred. No. 0.85;
Matches 91; Conservative 0; Mismatches 73; Indels 3; Gaps 1;

QY 47 TCCCCCAACTCCCIAGATGTCGTGGGGGCTGAACCTCGGCCGTTTAAAGGGCGGGC 106
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1157 TCCTCAAACTGCTGGGGTGGGGGAGGACTCATCTCATCCCATTA---GCCTGGC 1213
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 CCGGCTCCACGCTCTTCTGCTGAGTGAACATACATAACAGACGCGGAGACGGG 166
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1214 TGGGTATCCCTACCTTCTGGTCTCTGAGAGACTTACCTCTCCAGAGAGTCTCCGA 1273
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 GCGGGGAGGAGGAGACAGACAGCTTTGACCCATATACCTCTGCG 213
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1274 ATGAGTAGAAGGAGAGAGCTGATTGGAACITTAAGTACATCTGGG 1320

RESULT 3
US-09-122-126B-14/C
; Sequence 14, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(2910)
US-09-122-126B-14
```

```
Query Match: 9.3%; Score 32.6; DB 4; Length 3250;
Best Local Similarity 63.3%; Pred. No. 2.9;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 267 TCGAGCGAGAGTGAGTGGCGGCGGACCCGAGAGCGGACCTTCTCTCCCGG 326
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 317 TCGGCGAGGGGTGCGGGTGGCGGAGGCTCGGCTGCTGACCTCTCCCTGTC 258
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 CTGCGGCGAGGCGAGGCGG 345
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 CGCGGCGGGGCTGGGCGG 239

RESULT 4
US-09-741-154-3
; Sequence 3, Application US/09741154
; Patent No. 6437110
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CL001061
; CURRENT APPLICATION NUMBER: US/09/741,154
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16389
; TYPE: DNA
; ORGANISM: Human
US-09-741-154-3

Query Match: 9.1%; Score 32.2; DB 4; Length 16389;
Best Local Similarity 51.0%; Pred. No. 6;
```


Oy 121 C 121

Db 1299 C 1299

RESULT 8

US-09-750-560-1/C

: Sequence 1, Application US/09750580

: Patent No. 6455280

: GENERAL INFORMATION:

: APPLICANT: Yen, Frances

: APPLICANT: Denison, Blake

: APPLICANT: Bour, Barbara

: APPLICANT: Bihsain, Bernard

: APPLICANT: Dumas Milne Edwards, Jean-Baptiste

: APPLICANT: Duclert, Aymeric

: APPLICANT: Bouguetieret, Lydie

: APPLICANT: Ebbets-Reed, Dana

: APPLICANT: Satter-Cid, Luisa

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH

: FILE REFERENCE: 89 US2.CIP

: CURRENT APPLICATION NUMBER: US/09/750,560

: CURRENT FILING DATE: 2000-12-28

: PRIOR APPLICATION NUMBER: US 09/599,362

: PRIOR FILING DATE: 2000-06-21

: PRIOR APPLICATION NUMBER: PCT/IB00/0101

: PRIOR FILING DATE: 2000-06-21

: PRIOR APPLICATION NUMBER: PCT/IB99/02058

: PRIOR FILING DATE: 1999-12-20

: PRIOR APPLICATION NUMBER: US 49/469/099

: PRIOR FILING DATE: 1999-12-21

: PRIOR APPLICATION NUMBER: US 60/113,686

: PRIOR FILING DATE: 1998-12-22

: PRIOR APPLICATION NUMBER: US 60/141,032

: PRIOR FILING DATE: 1999-06-25

: NUMBER OF SEQ ID NOS: 6

: SOFTWARE: Patent.pm

: SEQ ID NO 1

: LENGTH: 81001

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

: NAME/KEY: misc.feature

: LOCATION: 10946..13946

: OTHER INFORMATION: 5'regulatory region

: NAME/KEY: exon

: LOCATION: 12947..13958

: OTHER INFORMATION: exon 1

: NAME/KEY: exon

: LOCATION: 13470..13526

: OTHER INFORMATION: exon 2

: NAME/KEY: exon

: LOCATION: 13641..13752

: OTHER INFORMATION: exon 3

: NAME/KEY: exon

: LOCATION: 14271..15968

: OTHER INFORMATION: exon 4

: NAME/KEY: misc.feature

: LOCATION: 15969..17969

: OTHER INFORMATION: 3'regulatory region

: NAME/KEY: allele

: LOCATION: 1239

: OTHER INFORMATION: 20-828-311 : polymorphic base C or T

: NAME/KEY: allele

: LOCATION: 12347

: OTHER INFORMATION: 17-42-319 : polymorphic base C or T

: NAME/KEY: allele

: LOCATION: 15241

: OTHER INFORMATION: 17-41-250 : polymorphic base C or T

: NAME/KEY: allele

: LOCATION: 42218

: OTHER INFORMATION: 20-841-149 : polymorphic base A or G

: NAME/KEY: allele

: LOCATION: 77039..77057

: LOCATION: 45442

: OTHER INFORMATION: 20-842-115 : polymorphic base A or G

: NAME/KEY: allele

: LOCATION: 77058

: OTHER INFORMATION: 20-853-415 : polymorphic base C or T

: NAME/KEY: primer_bind

: LOCATION: 929..949

: OTHER INFORMATION: 20-828.pu

: NAME/KEY: primer_bind

: LOCATION: 1357..1377

: OTHER INFORMATION: 20-828.rp complement

: NAME/KEY: primer_bind

: LOCATION: 12029..12050

: OTHER INFORMATION: 17-42.pu

: NAME/KEY: primer_bind

: LOCATION: 12581..12603

: OTHER INFORMATION: 17-42.rp complement

: NAME/KEY: primer_bind

: LOCATION: 14992..15012

: OTHER INFORMATION: 17-41.pu

: NAME/KEY: primer_bind

: LOCATION: 15480..15482

: OTHER INFORMATION: 17-41.rp complement

: NAME/KEY: primer_bind

: LOCATION: 45328..45347

: OTHER INFORMATION: 20-842.pu

: NAME/KEY: primer_bind

: LOCATION: 42070..42090

: OTHER INFORMATION: 20-841.pu

: NAME/KEY: primer_bind

: LOCATION: 42572..42591

: OTHER INFORMATION: 20-841.rp complement

: NAME/KEY: primer_bind

: LOCATION: 45328..45347

: OTHER INFORMATION: 20-842.pu

: NAME/KEY: primer_bind

: LOCATION: 45863..45883

: OTHER INFORMATION: 20-842.rp complement

: NAME/KEY: primer_bind

: LOCATION: 76644..76664

: OTHER INFORMATION: 20-853.pu

: NAME/KEY: primer_bind

: LOCATION: 77166..77185

: OTHER INFORMATION: 20-853.rp complement

: NAME/KEY: primer_bind

: LOCATION: 1220..1238

: OTHER INFORMATION: 20-828-311.mis

: NAME/KEY: primer_bind

: LOCATION: 1240..1258

: OTHER INFORMATION: 20-828-311.mis complement

: NAME/KEY: primer_bind

: LOCATION: 12328..12346

: OTHER INFORMATION: 17-42-319.mis

: NAME/KEY: primer_bind

: LOCATION: 12348..12366

: OTHER INFORMATION: 17-42-319.mis complement

: NAME/KEY: primer_bind

: LOCATION: 15222..15240

: OTHER INFORMATION: 17-41-250.mis

: NAME/KEY: primer_bind

: LOCATION: 15242..15260

: OTHER INFORMATION: 17-41-250.mis complement

: NAME/KEY: primer_bind

: LOCATION: 42199..42217

: OTHER INFORMATION: 20-841-149.mis

: NAME/KEY: primer_bind

: LOCATION: 42219..42237

: OTHER INFORMATION: 20-841-149.mis complement

: NAME/KEY: primer_bind

: LOCATION: 45423..45441

: OTHER INFORMATION: 20-842-115.mis

: NAME/KEY: primer_bind

: LOCATION: 45443..45461

: OTHER INFORMATION: 20-842-115.mis complement

: NAME/KEY: primer_bind

: LOCATION: 77039..77057

D5	611	AGAGACACCATGAGAACAGCATGTTGAGAGACATGGAGATGAGCGCAATGATCACATAC	552
QY	191	TTTACCGGATAGTAACCTCTGGCTGGTCTGACGGCGAATCTATAAAGGAACATAGTCCGG	250
D6	551	TTGGACTCATAGTAGATGATGAAACAATGTCGGCAGCTTATSCAGAAGTGTCACCAATGCAG	492
QY	251	GCAAAACCCGGTAAITGGCGACGAGATGACTGGGGCCGGACCCGAGAGCCGAGCCG	310
D6	491	GGATGATACCCCGCAGCGCTTCGCGGTATGATGTGTGTAGCGCAAGCATAGAAG	432
QY	311	ACCTTCTCTCCGGGCTCGGCA	334
D6	431	AIGTGATGTACCTATCCACCCCA	408

RESULT 5

US-09-097-231-17/c
; Sequence 17, Application US/09097231
; Patent No. 6278038

FILE NO. 0278058
; GENERAL INFORMATION:

APPLICANT: Cone, Roger D
Chen, Wenbiao

Low, Malcolm J

TITLE OF INVENTION: Mammalian Melanocortin Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM.

COM ON REMEMBER FORM,
; MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

CONSOLE: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

SOFTWARE: FALCONIN R
: CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,233

FILING DATE: 12-Jun-1998

CLASSIFICATION: <Unknown>

CLASSIFICATION: UNCLASSIFIED
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION.	
NAME: NO 6278038nao	Kevin F

NAME: NO. 0278036HALL, KEVIN L
REGISTRATION NUMBER: 35 303

REGISTRATOR NUMBER: 33,303
REFERENCE/DOCKET NUMBER: 96-386-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0

TELEPHONE: 312-313-0000
TELEFAX: 312-913-0002

TEL: 012-345678
FAX: 012-345678

INFORMATION FOR SEO ID NO: 17:

: : SEQUENCE CHARACTERISTICS: :
: : CONTINUATION FOR SEQ ID NO: 17: :

```

SEQUENCE CHARACTERISTICS:
      LENGTH: 978 base pairs

```

LENGTH: 5/8 Dase pair
TYPE: nucleic acid

TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE

STRANDEDNESS: linear
TOPOLOGY: linear

TOPOLOGY: LINEAR
MOLECULE TYPE: DNA (circular)

;	MOLECULE
:	FEATURE:

FEATURE:	NAME/REV.	DATE
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

NAME/KEY:	CDS
LOCATION:	1 075

LOCATION: I.975
SEQUENCE DESCRIPTION: SEQ ID NO. 17.

ISS-09-097-231-17

```

Query Match      8.9%; Score 31.2; DB 4; Length 978;
Best Local Similarity 47.1%; Pred. No. 5.2;
Matches 96; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

[illegible]

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      February 3, 2003, 14:33:17 ; Search time 19.8691 Seconds
              (without alignments)
              7959.555 Million cell updates/sec

Title:       US-09-536-14C-3_COPY_1292_1643
Perfect score: 352
Sequence:    1 ggaacctaaagacactgctg.....cagggcagggcgaggagcgc 352

Scoring table:  IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:     396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

ALIGNMENTS

RESULTS

RESOLUT I
US-09-846-456-1

: Sequence 1, Application US/09846456

; Patent No. US20020146792A1

; GENERAL INFORMATION:

; APPLICANT: ROSIER, Marie

; APPLICANT: Prades, Catherine

; APPLICANT: Lemcine, Cendrine

; APPLICANT: Naudin, Laurent

; APPLICANT: Denefle, Patrice

; APPLICANT: Duverger, Nicolas

APPLICANT: Brewer, Bryan

APPLICANT: Remaley, Alan

APPLICANT: FOJO, SILVIA

1. TITLE OF INVENTION: REGULATORY NUCLEIC ACID FOR THE ABC1 GENE, MOLECULES MODIFYING ACTIVITY AND CHARACTERISTICS OF THE ABC1 GENE, MOLECULES MODIFYING

1. TITLE OF INVENTION: ACTIVITY AND THERAPEUTIC USES OF

FILE REFERENCE: 3606.0303
: CURRENT APPLICATION NUMBER: US/99/846 456

; CURRENT AFFILIATION NUMBER: 03/0
; CURRENT FILING DATE: 2001-05-02

7. CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/207,280

PRIOR FILING DATE: 2000-05-02

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 3231

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-846-456-1

Query Match	95.3%	Score 335.4	DB 10	Length 3231
Best Local Similarity	98.3%	Pred. No. 1.5e+89		
Matches 351	Conservative 0	Mismatches 1	Indels 5	Gaps 1

QV 1 GGACCCCTAGACACCTGCTGTACCCCTCCAC-----CCCCACCCACCCACCTCCCCCAA 55

[illegible]

Db 2628 GGACCCCTAAGACACACCTGCTGTACCTCCACCCCAACCCCACTCCCCCA 26

[illegible]

QY 56 CTCCCAGATCTGTCGTGGCGGCTGAACGTCGCCCGTTAAGGGCGGCCCGGCTCC 11

Db 1359 CAGGTGGGCGCTGTCACAGCAGGTGATCTCGTCAGGTAGCGGGGACCCGAGTC 1300

QY 121 C 121

Db 1299 C 1299

RESULT 8

US-09-764-868-142

; Sequence 142, Application US/09764866
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patent-In Ver. 2.0
; SEQ ID NO 142
; LENGTH: 2928
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-142

Query Match 8.9%; Score 31.4; DB 9; Length 2928;

Best Local Similarity 56.2%; Pred. No. 6;

Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 GGACCTTAGACACCTGCTGTACCTCCACCCGCCACCCACCCACCTCCGCCCACTCC 60

Db 2464 GTACACTAGATCTGCTGCGGCTCCAGGCTGGACGACCCGATCCCACTCCCTCT 2523

QY 61 TAGATGTGCTGGGGGCTGACGTCGCCGCTTTAAGGGCGGG 105

Db 2524 GCTCTGATCTCGGGCTCAGCAGCCATCCCGCTGGGTGACGG 2568

RESULT 9

US-09-751-877-1/C

; Sequence 1, Application US/09751877

; Patent No. US20020142949A1

; GENERAL INFORMATION:

; APPLICANT: Yen, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bour, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Salter-Cid, Luisa
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH

; FILE REFERENCE: 89.US3.REG

; CURRENT APPLICATION NUMBER: US/09/751,877

; CURRENT FILING DATE: 2000-12-28

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Patent.pm

; SEQ ID NO 1

; LENGTH: 81001

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 10946..12946

; OTHER INFORMATION: 5'-regulatory region

; NAME/KEY: exon

; LOCATION: 12847..12958

; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 13470..13526
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 13641..13752
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 14271..15968
; OTHER INFORMATION: exon 4
; NAME/KEY: misc.feature
; LOCATION: 15969..17969
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1239
; OTHER INFORMATION: 20-828-311 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12347
; OTHER INFORMATION: 17-42-319 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15241
; OTHER INFORMATION: 17-41-250 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 42218
; OTHER INFORMATION: 20-841-149 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 45442
; OTHER INFORMATION: 20-842-115 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 77058
; OTHER INFORMATION: 20-853-415 : polymorphic base C or T
; NAME/KEY: primer_bind
; LOCATION: 929..949
; OTHER INFORMATION: 20-828.pu
; NAME/KEY: primer_bind
; LOCATION: 1357..1377
; OTHER INFORMATION: 20-828.rp complement
; NAME/KEY: primer_bind
; LOCATION: 12029..12050
; OTHER INFORMATION: 17-42.pu
; NAME/KEY: primer_bind
; LOCATION: 12581..12603
; OTHER INFORMATION: 17-42.rp complement
; NAME/KEY: primer_bind
; LOCATION: 14992..15012
; OTHER INFORMATION: 17-41.pu
; NAME/KEY: primer_bind
; LOCATION: 15460..15482
; OTHER INFORMATION: 17-41.rp complement
; NAME/KEY: primer_bind
; LOCATION: 42070..42090
; OTHER INFORMATION: 20-841.pu
; NAME/KEY: primer_bind
; LOCATION: 42572..42591
; OTHER INFORMATION: 20-841.rp complement
; NAME/KEY: primer_bind
; LOCATION: 45328..45347
; OTHER INFORMATION: 20-842.pu
; NAME/KEY: primer_bind
; LOCATION: 45863..45883
; OTHER INFORMATION: 20-842.rp complement
; NAME/KEY: primer_bind
; LOCATION: 76644..76664
; OTHER INFORMATION: 20-853.pu
; NAME/KEY: primer_bind
; LOCATION: 77166..77185
; OTHER INFORMATION: 20-853.rp complement
; NAME/KEY: primer_bind
; LOCATION: 1220..1238
; OTHER INFORMATION: 20-828-311.mis
; NAME/KEY: primer_bind
; LOCATION: 1240..1258
; OTHER INFORMATION: 20-828-311.mis complement


```
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (-)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match      8.58; Score 30.4; DB 10; Length: 99916;
Best Local Similarity 49.48; Pred. No. 38;
Matches 79; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 17 GCTGTACCTCCACCCACCCACCCACCTCCCTCCCAACTCCCTAGATGTTCGTGGGC 76
   || || || || || || || || || || || || || || || || || || ||
Db 3348 GCCGTCCTCTGTCGCACTCCCGACCTCCGCTGGCCGTGGGGGTGGAGGGCTGTGTG 3407

QY 77 GGTGGAAGTCGCCCGTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 136
   || || || || || || || || || || || || || || || || || || ||
Db 3408 GAGAGGATCACTGTGCTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3467

QY 137 TGAACATACATAACAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 176
   || || || || || || || || || || || || || || || || || || ||
Db 3468 TGTGTCAGAGAAATAGCACAGGTGTGAGGGGGGGGGGGGGGGGGGGGG 3507
```

Search completed: February 3, 2003, 16:33:24
Job time : 152.868 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: February 3, 2003, 16:39:45 ; Search time 263.012 Seconds
(without alignments)
8559.210 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1394_1532

Perfect score: 139

Sequence: 1 gggcccggtccacgtgct.....agccgaatataaaggaa 139

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estor:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_roc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	25.2	1863	17	AJ231825 Gallus ga
2	33.2	23.9	257	17	AL497552 T. Brucei
3	32.8	23.6	962	12	BG256511 602370051
4	32.4	23.3	712	13	BI890368
5	32.2	23.2	513	17	AQ209600 HS_3236-A
6	31.8	22.9	973	12	BG757993 602715037

RESULT 1

GGA200110

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

7	31.6	22.7	440	17	TA157G02P
8	31.6	22.7	516	17	AQ643297
9	31.6	22.7	742	13	BI193040
10	31.6	22.7	808	13	BI115921
11	31.4	22.6	782	17	CNS01YMS
12	31.2	22.4	437	17	BI103719
13	31.2	22.4	529	12	BG438512
14	31.2	22.4	875	13	BI133879
15	30.8	22.2	340	9	AA324852
16	30.8	22.2	846	17	AZ184185
17	30.8	22.2	1170	14	BM810841
18	30.6	22.0	886	9	AL547703
19	30.4	21.9	359	12	BG322077
20	30.4	21.9	872	12	BF254048
21	30	21.6	310	9	AL534701
22	30	21.6	528	12	BG081346
23	30	21.6	586	12	BF725798
24	30	21.6	611	13	BM011917
25	30	21.6	671	10	BE391123
26	30	21.6	677	10	BE272078
27	30	21.6	683	10	BE258315
28	30	21.6	693	10	SE389737
29	30	21.6	701	13	BI193161
30	30	21.6	705	14	BM991113
31	30	21.6	709	12	BG897930
32	30	21.6	718	13	BI561910
33	30	21.6	728	13	BI520520
34	30	21.6	771	9	AL556360
35	30	21.6	774	12	BG702611
36	30	21.6	827	12	BG821508
37	30	21.6	837	12	BG770188
38	30	21.6	849	14	BQ651632
39	30	21.6	870	14	BQ897942
40	30	21.6	892	12	BF312138
41	30	21.6	894	14	BQ670096
42	30	21.6	901	14	BQ431809
43	30	21.6	911	14	BQ222206
44	30	21.6	917	12	BG331708
45	30	21.6	922	13	BM451890

ALIGNMENTS

GGA200110 1863 bp DNA linear GSS 06-JUL-2000
Gallus gallus anonymous sequence from Cosmid mapping to a
microchromosome (Cosmid 20 - Contig 12), genomic survey sequence.

AJ231825

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Location/Qualifiers

Anim. Genet. 31 (2), 96-103 (2000)

20244064

10782207

2 (bases 1 to 1863)

Smith, J., Bruley, C.K., Paton, I.R., Law, A.S., Masabanda, J.,

Waddington, D., Fries, R. and Burt, D.W.

Direct Submission

Submitted (12-AUG-1998) Division of Molecular Biology, Roslin

Institute, Roslin, Midlothian EH25 9PS, Scotland, UK

```

source
1..1863
/organism="Gallus gallus"
/db_xref="taxon:9031"
/chromosome="microchromosome"
BASE COUNT      408 a   392 c   580 g   482 t
ORIGIN

Query Match      25.2%; Score 35; DB 17; Length 1863;
Best Local Similarity 55.3%; Pred. No. 4.5;
Matches 58; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 17 TCGTTTCCTGAGTGAAGTACTACATAACAGAGCGCGGAGGAGGAGG 76
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 TGATTACAGCAGAGTGAAGTACTACATAAGCAGCAGCAGTCTTTGGGGIAGC 298

QY 77 GAGACACAGCGTTTGACGGATAGTAACTCTGGCGCTCGTGACCGCAATCTATAAAG 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 299 GAAGAGAAATTCGAGCTCATTCAGGCACATCCACCCCTGGAGGTGAATTAACAG 358

QY 137 GAA 139
      |||
DB 359 GAA 361

RESULT 2
TA376C03P/c
LOCUS      TA376C03P      257 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION      T. brucei sheared genomic DNA clone 376c03, forward sequence,
                  genomic survey sequence.
ACCESSION      AL497552
VERSION      AL497552.1 GI:11873274
KEYWORDS      GSS.
SOURCE      Trypanosoma brucei.
ORGANISM      Trypanosoma brucei.
                  Eukaryota; Euklenozoa; Kinetoplastica; Trypanosomatidae;
                  Trypanosoma.
REFERENCE      1 (bases 1 to 257)
AUTHORS      Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
                  Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
                  Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE      Direct Submission
JOURNAL      Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
                  project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
                  Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
                  nh@sanger.ac.uk
COMMENT      Constructed at the Institute for Genomic Research (TIGR),
                  Rockville, MD. Genomic DNA isolated from a cloned population of
                  Trypanosoma brucei (TREF927/4 GUNat 10.1) was mechanically sheared
                  to give a tight size distribution (
                  4 kb). The v = 1 method used for the library construction is
                  described in detail in Smith, H. and Venter, J.C. (Making small
                  insert libraries for whole genome shotgun sequencing projects. In
                  Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
                  Barrell, Oxford University Press, 1999).
                  Email: nelsayed@tigr.org
                  Details of T. brucei sequencing at the Sanger Centre are available
                  at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES
source
1..257
Location/Qualifiers
BASE COUNT      74 a   72 c   54 g   57 t
ORIGIN

Query Match      23.9%; Score 33.2; DB 17; Length 257;
Best Local Similarity 57.8%; Pred. No. 9.9;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 28 GAGTGACTGAATACATAACAGCGCGGAGCGGCGGAGGAGGAGGACACAC 87
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 GGGTGCTGCATCCGATGCTGTACGCCGAGTGAGAGGGGGGGAACGACTACTGC 198

```

```

QY 88 CTTTGACCGATAGTAACCTCTGGCTGGTGCAGCGGAATCT 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 GTTTGACCGCTCCTAACCACTAGTTTGAATGTACCAACTCT 156

RESULT 3
BG256511      962 bp      mRNA      linear      EST 13-FEB-2001
LOCUS      602370051F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4478196 5',
DEFINITION      mRNA sequence.
ACCESSION      BG256511
VERSION      BG256511.1 GI:12766327
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 962)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
                  Email: ccspbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM10308 row: e column: 13
                  High quality sequence start: 2
                  High quality sequence stop: 733.
FEATURES
source
1..962
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4478196"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIH_MGC Library."
BASE COUNT      193 a   276 c   314 g   178 t
ORIGIN

Query Match      23.6%; Score 32.8; DB 12; Length 962;
Best Local Similarity 67.6%; Pred. No. 18;
Matches 46; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 16 GTGCTTTCGTGCTGAGTACTAGACTACATAACAGAGCGCGGAGGAGGAGGAG 75
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 780 GTACTTCCTGTGCTGAGTACTAGAGAAACCGTCCAGGGGGGCGCGGAGGAGGTG 839

QY 76 GGAGAGCA 83
      ||| |||
DB 840 GGACACA 847

RESULT 4
BI890368/c
LOCUS      BI890368      712 bp      mRNA      linear      EST 12-OCT-2001
DEFINITION      ZF637-2-032123 zebrafish shield stage whole embryo cDNA library
                  MPMGP637 Danio rerio cDNA clone MPMGP637_7P1; MPMGP637P017 5', mRNA
                  sequence.
ACCESSION      BI890368
VERSION      BI890368.1 GI:16097639
KEYWORDS      EST.
SOURCE      zebrafish.

```



```

Tel: (206) 616-3678
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3236 row: E column: 20
Class: BAC ends
High quality sequence stop: 513.
Location/Qualifiers
1. .513
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=3236 Col=20 Row=E"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 108 a 118 c 101 g 175 t 11 others
ORIGIN
Query Match 23.2%; Score 32.2; DB 17; Length 513;
Best Local Similarity 61.2%; Pred.No.23;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 36 GAACGTCATACAGAGCCGGACCGCGCGGAGGAGGAGGACAGCGCTTTGACC 95
|||||
Db 186 GAACACATACACATGTGCGCTCTATGGGGTGGGGGAGGAGGACATCAGGAAGATA 127
|||||
QY 96 GATAGTAACTCTCGCTCGGTGCA 120
|||||
Db 126 GATAAAATCTGGGTAGATACA 102
|||||
RESULT 6
BG757993/c

```

```

mRNA sequence.
BG757993.1 GI:14068646
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 973)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: ecgaps-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI704 row: 1 column: 09
High quality sequence stop: 498.
Location/Qualifiers
1..973
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4855376"
/clone_lib="NIH-MGC_48"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by cigo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of

```



```

ACCESSION      BI193040
VERSION        BI193040.1  GI:14648060
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 742)
AUTHORS        NIH-MGC http://mgi.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: ATCC
               CDNA Library Preparation: Ling Hong/Rubin Laboratory
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone Distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1CMI852  row: f  column: 23
               High quality sequence start: 15
               High quality sequence stop: 393.
               Location/Qualifiers
                 1..742
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone_lib="IMAGE:5090326"
                 /clone_lib="NIH_MGC_42"
                 /tissue_type="epithelioid carcinoma cell line"
                 /lab_host="DH10B (phage-resistant)"
                 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                 Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally
                 cloned into EcoRI/XhoI sites using the following 5'
                 adaptor: GCCACGAG(G). Size-selected >500bp for average
                 for average insert size 1.8kb. Library constructed by Ling
                 Hong in the laboratory of Gerald M. Rubin (University of
                 California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and Superscript II RT (Life Technologies).
                 Note: this is a NIH_MGC Library. |"
BASE COUNT    175 a 240 c 197 g 130 t
ORIGIN
Query Match   22.7%  Score 31.6;  DB 13;  Length 742;
Best Local Similarity 62.8%;  Pred. No. 38;
Matches       49;  Conservative 0;  Mismatches 29;  Indels 0;  Gaps 0;

QY  1  GGGCCCGGCTCCACGTCCTTCGTGAGTGACTGACTACATAACAGAGCGCGGAA 60
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
Db  151 GGGCTGCTGCTCCGTCGTCCTCGGCGCTGATGGCGGCAACAGGTGATCAGGGGCGAGCGGA 92
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
QY  61  CGGGCGCGGGGAGGAGGA 78
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
Db  91  CGGGCGCGGGGCGGCGGA 74
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1

RESULT 10
BI115921/c
LOCUS         BI115921      808 bp  mRNA  linear  EST 26-JUN-2001
DEFINITION   602866693F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5015138 5',
              mRNA sequence.
ACCESSION    BI115921
VERSION      BI115921.1  GI:14566822
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 808)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CMI823  row: j  column: 23
High quality sequence stop: 500.
Location/Qualifiers
  1..808
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="IMAGE:5015138"
  /clone_lib="NIH_MGC_7"
  /tissue_type="small cell carcinoma"
  /cell_line="MGC3"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
  EcoRI; cDNA made by oligo-dt priming. Directionally
  cloned into EcoRI/XhoI sites using the following 5'
  adaptor: GCCACGAG(G). Size-selected >500bp for average
  insert size 1.8kb. Library constructed by Ling Hong in
  the laboratory of Gerald M. Rubin (University of
  California, Berkeley) using ZAP-cDNA synthesis kit
  (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT    201 a 237 c 227 g 143 t
ORIGIN
Query Match   22.7%  Score 31.6;  DB 13;  Length 808;
Best Local Similarity 52.8%;  Pred. No. 39;
Matches       49;  Conservative 0;  Mismatches 29;  Indels 0;  Gaps 0;

QY  1  GGGCCCGGCTCCACGTCCTTCGTGAGTGACTGACTACATAACAGAGCGCGGAA 60
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
Db  150 GGGCTGCTGCTCCGTCGTCCTCGGCGCTGATGGCGGCAACAGGTGATCAGGGGCGAGCGG 91
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
QY  61  CGGGCGCGGGGAGGAGGA 78
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1
Db  90  CGGGCGCGGGGCGGCGGA 73
      ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1

RESULT 11
CNS01YK5/c
LOCUS         CNS01YK5      782 bp  DNA  linear  GSS 12-MAY-2000
DEFINITION   Tetraodon nigroviridis genome survey sequence I7 end of clone
              218D02 of library G from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION    AL173210
VERSION      AL173210.1  GI:7811267
KEYWORDS     GSS: genome survey sequence.
SOURCE       Tetraodon nigroviridis.
ORGANISM     Tetraodon nigroviridis
REFERENCE    1 (bases 1 to 782)
AUTHORS      Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
              Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
              Saurin,W. and Weissenbach,J.
              Human gene number estimate provided by genome wide analysis using
              Tetraodon nigroviridis DNA sequence
              Unpublished
              2 (bases 1 to 782)
              Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
              Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
              Weissenbach,J.
              Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish Tetraodon nigroviridis
              Unpublished
              3 (bases 1 to 782)
              JOURNAL
              REFERENCE
              AUTHORS
              TITLE
              JOURNAL
              REFERENCE
              AUTHORS
              Genoscope.
```

RCTI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pRAREcl cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA.

BASE COUNT	100 a	111 c	106 g	120 t
ORIGIN				

	Query Match	22.4%	Score 31.2	DB 17	Length 437
	Best Local Similarity	63.2%	Pred. No. 44		
	Matches 48	Conservative 0	Mismatches 28	Indels 0	Gaps 0
QY	27	TGAGTGACTGAACATACATAACAGAGGCGCGGGAACGGCGGGGAGGAGGAGGCACAG	86		
Db	201	TGGTACTAANAAGTACAGGCGCTCGAGGCGCGGAGGGGGGGGGGAGAGGCGCAG	142		

67 GCCTTACCCGATGAGA 102
 ||| |
 141 TTCAATACGTAAAGTA 126

RESULT 13	LOCUS	DEFINITION	EST 10-MAY-2001
EG438512/c	EG438512	529 bp mRNA	linear
	ps27604.y2	Trichinella spiralis ML CMVSPORT	Jasmer
	spiralis cDNA 5'	similar to SW:CG42 DROME P40793	CD442 HOMOLOG [5]
	TR:Q9U465	TR:Q9U956	TR:Q9U954 ; mRNA sequence.
	EG438512		
	ACCESSION		

RECORDS	SOURCE	ORGANISM	REFERENCE
251.		<i>Trichinella spiralis</i> .	
		<i>Trichinella spiralis</i>	
		Eukaryota; Metazoa; Nematoda; Enopleia; Trichocephalida;	
		Trichinellidae; <i>Trichinella</i> .	
		1 (bases: 1 to 529)	

McCartee, J., Carlson, S., Chaperlin, B., Fape, D., Mathis, D., Wyler, R., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, J., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Pearson, D., Swallier, T., Harvey, N., Schmitt, P., Kohn, S.

TITLE	JOURNAL	COMMENT
Wilson, R. The Washington Univ. Nematode ESI Project, 1999 Unpublished (1999)		Contact: McCarter, JP Nematode ESI Project, 1999
, Shatt, T., Jackson, V., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.		

the Washington Univ. Nematode EST Project, 1999.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Dr. Doug Jasmer
(djasmer@wemmed.wsu.edu) at Washington State University, Dept. of
Veterinary Microbiology and Pathology DNA Sequencing by: Washington
University Genome Sequencing Center, St. Louis

```

FEATURES
  source
    1..529
      /organism="Trichinella spiralis"
      /db_xref="taxon:6334"
      /clone_lib="Trichinella spiralis ML CMVSPORT.tasmer"
      seq_primer: Spc
      High quality sequence stop: 422.
      Location/Qualifiers

```

/lab_host="DH-08"
/note="vector: pcwmsport-7.neo: Site_1: NotI; Site_2: SalI
; The library was constructed using mRNA isolated from
total RNA with oligo-dT cellulose. Total RNA was generated
from muscle larvae that were isolated from infected rats.
Larvae were liberated by pepsin/HCl digestion, incubated
with 8 SDS, treated with RNase and DNase to eliminate
host nucleic acid contamination, and purified on a Percoll
gradient. The T. spiralis isolate was obtained from
Dickson Despotnikar (Columbia University). The library was
provided by Dr. Doug Jasmer (diagram@vetmed.wsu.edu) and

colleagues at Washington State University, DNA Sequencing
by: Washington University Genome Sequencing Center St.
Louis."

BASE COUNT 161 a 107 c 119 g 142 t
ORIGIN

Query Match 22.4%; Score 31.2; DB 12: Length 529;
Best Local Similarity 54.3%; Pred. No. 46;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 11 TCCACGCTTCTGCTGAGTACTACTATAAACAAGCGCGGAGCGGCGGG 70

DB 228 TTCACATTTTCATAGCTGGATGCGCAAACTACAGAAAACAAGAGAGATCCGTC 169

QY 71 AGGAGGGAGACACAGGCTTTGACCGAGTAGTAACCTTCGCTCGGTCCAGCCGAA 126

DB 168 TGGGATAGGACAGTGGCTCAACAGCATCGTAATCTTCTGTCCGCGAGTGCAGAA 113

RESULT 14
BI193879
LOCUS 502948207F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5091625 5',
DEFINITION mRNA sequence.
ACCESSION BI193879
VERSION BI193879.1 GI:14648899
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs.r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LICM1855 row: m column: 02
High quality sequence stop: 738.
Location/Qualifiers
1..875

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5091625"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library. |"

BASE COUNT 176 a 290 c 222 g 187 t
ORIGIN
Query Match 22.4%; Score 31.2; DB 13: Length 875;
Best Local Similarity 53.2%; Pred. No. 52;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 5 CCGCGCTCACAGTGTCTTCTGCTGAGTACTGAATACATAAACAAGCGCGGACGGG 64

DB 81 CCCCACCCACGCTCTGTGGACCCCGGA-TATTTACGCTTCAGACAGACGCGGGG 140

QY 65 GCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACCTTCGCGTCCGGTGCAGCG 124
DB 141 GCAGAGGGGCCCGCGGAGGGGTGCTACTTCTTAAACACCTTCGCGGCTGCTTAGTCA 200

QY 125 AATC 128

DB 201 CAGC 204

RESULT 15
AA324852
LOCUS 340 bp mRNA linear EST 20-APR-1997
DEFINITION EST27749 Cerebellum II Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA324852
VERSION AA324852.1 GI:1977096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 340)
AUTHORS Adams.M.D., Kerlavage.A.R., Fleischmann.R.D., Fuldner.R.A., Bult
C.J., Lee.N.H., Kirkness.E.F., Weinstock.K.G., Gocayne.J.D., White
O., Sutton.G., Blake.J.A., Brandon.R.C., Man-Wai.C., Clayton.R.A.,
Cline.T.R., Cotton.M.D., Earle-Hughes.J., Fine.L.D., Fitzgerald
L.M., Fitzhugh.W.M., Fritchman.J.L., Geoghegan.N.S., Glodek.A.,
Gnehm.C.L., Hanna.M.C., Hedblom.E., Hinkley.P.S., Kelley.J.M.,
Kelley.J.C., Liu.L.-I., Marmaros.S.M., Merrick.J.M.,
Moreno-Palauques.R.F., McDonald.L.A., Nguyen.D.T., Pelligrino.S.M.,
Phillips.C.A., Ryder.S.E., Scott.J.L., Saudek.D.M., Shirley.R.,
Small.K.V., Spriggs.T.A., Utterback.T.R., Weidman.J.F., Li.Y.,
Bednarek.D.P., Cao.L., Cepeda.M.A., Coleman.J.A., Collins.E.J.,
Dinkeld., Feng.D.-F., Ferrie.A., Fischer.C., Hastings.G.A., He.W.W.,
Kunsch.C., Hungjun.J., Li.H., Weissner.P.S., Olsen.H., Raymond.L.,
Wei.Y.F., Wing.J., Xu.C., Yu.G.L., Ruben.S.M., Dillion.P.J., Fannon
M.R., Rosen.C.A., Haseltine.W.A., Fields.C., Fraser.C.M. and
Venter.J.C.

Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tadb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1..340
/organism="Homo sapiens"
/db_xref="ATCC (inhost):125343"
/db_xref="taxon:9606"
/clone_lib="Cerebellum II"
/tissue_type="cerebellum"
/dev_stage="adult"
/note="Organ: brain; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI"
57 a 129 c 70 t 6 others

BASE COUNT 57 a 129 c 70 t 6 others
ORIGIN

Query Match 22.2%; Score 30.8; DB 9: Length 340;
Best Local Similarity 52.1%; Pred. No. 55;
Matches 62; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 5 CCGCGCTCACAGTGTCTTCTGCTGAGTACTGAATACATAAACAAGCGCGGACGGG 64

DB 81 CCCCACCCACGCTCTGTGGACCCCGGA-TATTTACGCTTCAGACAGACGCGGGG 140

Db 163 CCCATNTCCGGGAGAGCCCGGAGACCCCTGAACTACTTCTCGGTGCTGGCGGGAG 222
QY 65 GCGGGGAGGGGAGAGCACAGGCTTTGACCGATAGTAACTCTGGGCTCGGTGCAGCC 123
Db 223 GCTGRANTTGGGAGCAGCAGANAGTTGGCCCTGCCACCCCAACCTCTCTCCAGCC 281

Search completed: February 4, 2003, 03:27:17
Job time : 277.726 secs


```

; NAME/KEY: primer_bind
; LOCATION: 77039..77057
; OTHER INFORMATION: 20-853-415.mis
; NAME/KEY: primer_bind
; LOCATION: 77059..77077
; OTHER INFORMATION: 20-853-415.mis complement
; NAME/KEY: misc_binding
; LOCATION: 227..1251
; OTHER INFORMATION: 20-828-311.probe
; NAME/KEY: misc_binding
; LOCATION: 12335..12359
; OTHER INFORMATION: 17-42-319.probe
; NAME/KEY: misc_binding
; LOCATION: 15229..15253
; OTHER INFORMATION: 17-41-250.probe
; NAME/KEY: misc_binding
; LOCATION: 42206..42230
; OTHER INFORMATION: 20-941-149.probe
; NAME/KEY: misc_binding
; LOCATION: 45430..45454
; OTHER INFORMATION: 20-842-115.probe
; NAME/KEY: misc_binding
; LOCATION: 77046..77070
; OTHER INFORMATION: 20-853-415.probe
; US-09-750-580-1

Query Match          21.6%; Score 30; DB 4; Length 81001;
Best Local Similarity 61.5%; Pred. No. 1.9;
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 GGGCCCGCGGCCAGTCTTCTGCTGAGTGAACATACATAAACAGAGCGCGGAA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 17479 GGGCTGCTGCTCTCTGCTGCGGCTGATCGGCCGGAACAGGTGATCAGGGCGAGCGG 17420

QY 61 GGGGGCGGGGAGAGGGA 78
    ||||| ||||| |||||
DQ 17419 GGGGGCGGGGCGGGA 17402

RESULT 3
US-09-276-531-56
; Sequence 56, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandmar, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998

```

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murty, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1610 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUP08
; CLONE: 1395833
; US-09-276-531-56

Query Match          20.4%; Score 28.4; DB 4; Length 1610;
Best Local Similarity 53.6%; Pred. No. 1.7;
Matches 59; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 2 GGGCCCGGCTCCAGTGTCTTCTGCTGAGTGAACATACATAAACAGAGCGCGGAAAC 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1211 GGCTGGGTATTCCTACTTCTGCTGCTGAGAGACTTACTCTTCCAGAGATCTCTCC 1270

QY 62 GGGCGGGGAGGAGGAGACAGAGCGTTTGACCGATAGTACCTCTCGG 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1271 CGAATGATAGAGGAGAGAGCTGGATTTGAACITTAAGTGACATCTGGG 1320

RESULT 4
US-09-212-247C-3/c
; Sequence 3, Application US/09212247C
; Patent No. 6391603
; GENERAL INFORMATION:
; APPLICANT: POMPEJUS, Markus; SUTELBERGER, Harald; JOEFFKEN, Hans
; and GARCIA, Maria Angeles Santos
; WOLFGANG DOVAL, Jose Luis Revuelta; JIMENEZ, Alberto;
; TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii
; and the use thereof in microbial riboflavin
; synthesis
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM AT-compatible, Pentium processor
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,247C
; FILING DATE: 16-Dec-1998
; CLASSIFICATION: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5369 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..54
; FEATURE:

```



```

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-68

Query Match
Best Local Similarity 19.3%; Score 26.8; DB 3; Length 966;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY 3 GCCCGGCTCCAGGCTTCTGCTGAGTGACTGAACATACATAAAGAGAGCGGGAACG 62
DB 522 GCCCTGGATCTTCCTCACTTGTCTAGACACITGACGAGAGCTCTGGGCGAGGAGCT 463

QY 63 GGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCCCTCGGTGCAGC 122
DB 462 GGCAGGGCCCTAATGGTGTAGCCATGTTGAGATTTATGAGATTCCTTTAGACGGAGGAGA 403

QY 123 CGAATCTATAAAG 136
DB 402 CGGGTTGATAGTAG 389

RESULT 8
US-08-469-318-70/c
; Sequence 70, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-C8-469-318-70

Query Match
Best Local Similarity 19.3%; Score 26.8; DB 3; Length 966;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY 3 GCCCGGCTCCAGGCTTCTGCTGAGTGACTGAACATACATAAAGAGAGCGGGAACG 62
DB 522 GCCCTGGATCTTCCTCACTTGTCTAGACACITGACGAGAGCTCTGGGCGAGGAGCT 463

QY 63 GGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCCCTCGGTGCAGC 122
DB 462 GGCAGGGCCCTAATGGTGTAGCCATGTTGAGATTTATGAGATTCCTTTAGACGGAGGAGA 403

QY 123 CGAATCTATAAAG 136
DB 402 CGGGTTGATAGTAG 389

RESULT 9
US-08-469-318-71/c
; Sequence 71, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-71

Query Match
Best Local Similarity 19.3%; Score 26.8; DB 3; Length 966;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY 3 GCCCGGCTCCAGGCTTCTGCTGAGTGACTGAACATACATAAAGAGAGCGGGAACG 62
DB 522 GCCCTGGATCTTCCTCACTTGTCTAGACACITGACGAGAGCTCTGGGCGAGGAGCT 463

QY 63 GGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCCCTCGGTGCAGC 122
DB 462 GGCAGGGCCCTAATGGTGTAGCCATGTTGAGATTTATGAGATTCCTTTAGACGGAGGAGA 403

QY 123 CGAATCTATAAAG 136
DB 402 CGGGTTGATAGTAG 389

RESULT 10
US-08-468-609A-68/c
; Sequence 68, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Adams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Bradford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maïre H.
```

```

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-71

Query Match
Best Local Similarity 19.3%; Score 26.8; DB 3; Length 966;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY 3 GCCCGGCTCCAGGCTTCTGCTGAGTGACTGAACATACATAAAGAGAGCGGGAACG 62
DB 522 GCCCTGGATCTTCCTCACTTGTCTAGACACITGACGAGAGCTCTGGGCGAGGAGCT 463

QY 63 GGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCCCTCGGTGCAGC 122
DB 462 GGCAGGGCCCTAATGGTGTAGCCATGTTGAGATTTATGAGATTCCTTTAGACGGAGGAGA 403

QY 123 CGAATCTATAAAG 136
DB 402 CGGGTTGATAGTAG 389

RESULT 9
US-08-469-318-71/c
; Sequence 71, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-71

Query Match
Best Local Similarity 19.3%; Score 26.8; DB 3; Length 966;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY 3 GCCCGGCTCCAGGCTTCTGCTGAGTGACTGAACATACATAAAGAGAGCGGGAACG 62
DB 522 GCCCTGGATCTTCCTCACTTGTCTAGACACITGACGAGAGCTCTGGGCGAGGAGCT 463

QY 63 GGGCGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCCCTCGGTGCAGC 122
DB 462 GGCAGGGCCCTAATGGTGTAGCCATGTTGAGATTTATGAGATTCCTTTAGACGGAGGAGA 403

QY 123 CGAATCTATAAAG 136
DB 402 CGGGTTGATAGTAG 389

RESULT 10
US-08-468-609A-68/c
; Sequence 68, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Adams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Bradford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maïre H.
```



```

; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (IL-
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-609A-71

Query Match          19.3%; Score 26.8; DB 3; Length 966;
Best Local Similarity 50.0%; Pred. No. 5.1;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY  3  GCCCGGCTCCAGTGTCTTCTGAGTCACTGAACTAATAAAGAGCGCGGGAACG 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  522  GCCCTGGAATCTCTCACTGCTTACGACCTTACGAGGAGCTCTGGGGCAGGAGCT 463

QY  63  GGGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  462  GGCAGGGCCCTAATGCTGAGCCATGTTGGAGATTATGAGATTCTTTAGACGGAGGAGA 403

QY  123  CGAATCTATAAAG 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  402  CCGGTTGATAGTAG 389

RESULT 13
US-08-446-872A-68/c
; Sequence 68, Application US/08446872A
; Patent No. 6361977
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maire H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Orlins, Peter O.
; APPLICANT: Paik, Kumhan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,872A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-446-872A-68

Query Match          19.3%; Score 26.8; DB 4; Length 966;
Best Local Similarity 50.0%; Pred. No. 5.1;
Matches 67; Conservative 0; Mismatches 67; Indels 0; Gaps 0:

QY  3  GCCCGGCTCCAGTGTCTTCTGAGTCACTGAACTAATAAAGAGGCGGGAACG 52
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  522  GCCCTGGAATCTCTCACTGCTTACGACCTTACGAGGAGCTCTGGGGCAGGAGCT 463

QY  63  GGGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  462  GGCAGGGCCCTAATGCTGAGCCATGTTGGAGATTATGAGATTCTTTAGACGGAGGAGA 403

QY  123  CGAATCTATAAAG 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  402  CCGGTTGATAGTAG 389

RESULT 14
US-08-446-872A-70/c
; Sequence 70, Application US/08446872A
; Patent No. 6361977
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maire H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Orlins, Peter O.
; APPLICANT: Paik, Kumhan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,

```



```

Query Match          20.6%; Score 28.6; DB 9; Length 975;
Best Local Similarity 52.0%; Pred. No. 2.1;
Matches 64; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCGCGCTCCAGTCTTCTGCTGAGTGAACATACATAACAGAGCGCGGGAACGG 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CTCACACCTCTCTCTGCTCCAGCAGGTAAACACTCAATCCGGTGGCTGTAGGAG 623

QY 64 GCGGGGAGGAGGAGGAGACAGAGCTTTGACCATAGTAACTCTCGGTTCGTGAGGC 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 GAGCGGAGGAGGAGGAGGAGGATCCTGAGAGAGATAGAACTTCGAACTCCCTGGAAGAC 563

QY 124 GAA 126
    ||
Db 562 TAA 560

RESULT 6
US-09-294-093B-6019
; Sequence 6019, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalquidi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 6019
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700382861H1
; NAME/KEY: unsure
; LOCATION: 279
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-6019

Query Match          19.9%; Score 27.6; DB 10; Length 280;
Best Local Similarity 63.8%; Pred. No. 3.1;
Matches 42; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 21 TTCTGCTGAGTGAATGAATACATAAAGAGGCGCGGGAACGGGCGGAGGAGGAGA 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 TCCGCTGAGATCTGACACAGATGGACAGCAGGACGAGAGAGGCTGCTGAGGGTGGCGA 198

QY 81 GCACAG 86
    |||
Db 199 TGACGG 204

RESULT 7
US-09-796-692-6574
; Sequence 6574, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01

```

```

; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-03-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6574
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (341)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (362)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (383)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-6574

Query Match          19.7%; Score 27.4; DB 9; Length 393;
Best Local Similarity 53.2%; Pred. No. 4;
Matches 58; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 2 GGGCCCCGCTCCAGCTCTCTTCTGCTGAGTGAACATACATAAAGAGGCGCGGGAAC 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GCGCCCGGATGCCCGCAGCAGTGTCTCTCAGAGTGGAGCTGCAAGACCTGAGGAAGAGCT 240

QY 62 GGGCGGGGAGGAGGAGGAGCAGCAGGCTTTACCGCATAGTAACCTCTGC 110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GGAGAGAGTGCAGGAGAGGCTCCAGGCTCAGGTGGGAAAACATCAGTGC 289

RESULT 8
US-10-044-090-30/c
; Sequence 30, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 1466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137881A1 4843169CB1
US-10-044-090-30

Query Match      19.7%; Score 27.4; DB 12; Length 1466;
Best Local Similarity 59.7%; Pred. No. 6.1;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 63 GGGCGGGGAGGAGGAGACAGCGCTTTCGTGAGTACTAGTAACTCTGGCGTGGTGCAGC 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 98 GGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 39
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 123 CGAATCTATATAAGAA 139
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 38 CTTAAAAAAGAAAAA 22
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-09-880-107-3768
; Sequence 3768, Application US/09860107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3768
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(66109)
; OTHER INFORMATION: n = a o c o z g o r t
US-09-880-107-3768

Query Match      19.6%; Score 27.2; DB 10; Length 66109;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 GGGCCCCGGCTCCAGCTGCTTCTGTGAGTACTAGTAACTATACATAAACACAGAGCCGGGAA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 34911 GGGTCGGGGTAGCCTTCAGTCCCGGAGAGCGCCAGACCCAAAGAGAGGCCACATGGGG 34970
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 34971 ATGGGGCTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 35030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 GCCGAATCTATAAAG 136
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 35031 CCGAAGGAGAGCAAG 35046
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-09-880-107-3814
; Sequence 3814, Application US/09860107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814

Query Match      19.6%; Score 27.2; DB 10; Length 198285;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 68; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1 GGGCCCCGGCTCCAGCTGCTTCTGTGAGTACTAGTAACTATACATAAACACAGAGCCGGGAA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1-0700 GGGTCGGGGTAGCCTTCAGTCCCGGAGAGCGCCAGACCCAAAGAGAGGCCACATGGGG 110759
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 110760 ATGGGGCTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 110819
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 GCGGAATCTATAAAG 136
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 110820 CCGAAGGAGAGCAAG 110835
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-08-900-220C-7
; Sequence 7, Application US/08900220C
; Patent No. US20020045206A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Ningning
; APPLICANT: Wang, Monica
; APPLICANT: Mahanthappa, Nagesh K.
; APPLICANT: Pang, Kevin
; APPLICANT: Jin, Ping
; TITLE OF INVENTION: Method of Treating Dopaminergic and
; GABA-nergic Disorders
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSER: POLEY, HOAG & ELIOT LLP
; STREET: ONE POST OFFICE SQUARE
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,220C
; FILING DATE: 24-Jul-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: ONV-044.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```
/
/ LENGTH: 1622 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: both
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 51..1283
/ SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-900-220C-7

Query Match
Best Local Similarity 19.4%; Score 27; DB 8; Length 1622;
Matches 63; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 16 GTGCTTCTCTGCTGAGTCACTCACTAATACAGAGCGCGGGAACGGGGCGGGGAGGAG 75
Db 1226 GCGTCTCTCTGCTAGAGAGGCGACGCTCCACCCACTGGGATGTCGGGGCGGAGGAGCTG 1285

QY 76 GGAGAGCAGCAGCGCTTTCAGCCGATAGTAACTCTCGGCTCGGTCGAGCGCGGATCTATAAA 135
Db 1286 AAAGGACTCCACCGCTGCGCTCTCTGGACTCTGTACTGGTCCAGAGCGCTCTCAGCCA 1345

QY 136 GGA 138
Db 1346 GGA 1348

RESULT 12
US-09-151-999-7
/ Sequence 7, Application US/09151999
/ Patent No. US20020151460A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Elizabeth
/ TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
/ FILE REFERENCE: CNV-031-02
/ CURRENT APPLICATION NUMBER: US/09/151.999
/ CURRENT FILING DATE: 1998-08-11
/ EARLIER APPLICATION NUMBER: 08/955,552
/ EARLIER FILING DATE: 1997-10-20
/ NUMBER OF SEQ ID NOS: 28
/ SOFT-WARE: Patentin ver. 2.0
/ SEQ ID NO 7
/ LENGTH: 1622
/ TYPE: DNA
/ ORGANISM: Homo sapien inh
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (51)..(1283)
US-09-151-999-7

Query Match
Best Local Similarity 19.4%; Score 27; DB 10; Length 1622;
Matches 63; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 16 GTGCTTCTCTGCTGAGTCACTCACTAATACAGAGCGCGGGAACGGGGCGGGGAGGAG 75
Db 1226 GCGTCTCTCTGCTAGAGAGGCGACGCTCCACCCACTGGGATGTCGGGGCGGAGGAGCTG 1285

QY 76 GGAGAGCAGCAGCGCTTTCAGCCGATAGTAACTCTCGGCTCGGTCGAGCGCGGATCTATAAA 135
Db 1286 AAAGGACTCCACCGCTGCGCTCTCTGGACTCTGTACTGGTCCAGAGCGCTCTCAGCCA 1345

QY 136 GGA 138
Db 1346 GGA 1348

RESULT 13
US-09-874-503-17
/ Sequence 17, Application US/09874503
/ Patent No. US20020177188A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Jian
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Li, Hanzhong
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Hymowitz, Sarah G.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Starovasnik, Melissa A.
/ APPLICANT: VanLoenen, Menno
/ APPLICANT: Vandlen, Richard
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Woog, William I.
/ APPLICANT: Yansura, Daniel
/ TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES THEREOF
/ FILE REFERENCE: PL381R1C1P3(US)
/ CURRENT APPLICATION NUMBER: US/09/874,503
/ CURRENT FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US 60/253,646
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: US 60/244,072
/ PRIOR FILING DATE: 2000-10-26
/ PRIOR APPLICATION NUMBER: US 60/242,837
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/175,481
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: US 60/191,007
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: US 60/213,807
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: US 60/172,096
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: US 60/138,387
/ PRIOR FILING DATE: 1999-06-09
/ PRIOR APPLICATION NUMBER: US 60/134,287
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: US 60/131,022
/ PRIOR FILING DATE: 1999-04-26
/ PRIOR APPLICATION NUMBER: US 60/130,232
/ PRIOR FILING DATE: 1999-04-21
/ PRIOR APPLICATION NUMBER: US 60/113,621
/ PRIOR FILING DATE: 1998-12-23
/ PRIOR APPLICATION NUMBER: US 60/085,579
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: US 09/854,208
/ PRIOR FILING DATE: 2001-05-10
/ PRIOR APPLICATION NUMBER: US 09/854,280
/ PRIOR FILING DATE: 2001-05-20
/ PRIOR APPLICATION NUMBER: US 09/816,744
/ PRIOR FILING DATE: 2001-03-22
/ PRIOR APPLICATION NUMBER: US 09/747,259
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: US 09/644,848
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: US 09/380,142
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: US 09/380,138
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: US 09/311,832
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: US PCT/US01/06520
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US PCT/US00/34956
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: US PCT/US00/32678
/ PRIOR FILING DATE: 2000-12-01
/ PRIOR APPLICATION NUMBER: US PCT/US00/30873
/ PRIOR FILING DATE: 2000-11-10
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 : Search time 715.76 seconds
(without alignments)
10165.006 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1394_1643

Perfect score: 250

Sequence: 1 gggccggcgctcccgctgct.....caggcgagggcgaggagtc 250

Scoring table: IDENT-1Y.NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.pl.*
- 8: gb.pr.*
- 9: gb.ro.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	250	100.0	1167	9	HS252201	Homo sapi
2	250	100.0	1643	6	AX060715	Sequence
3	250	100.0	1643	6	AX060894	Sequence
4	250	100.0	69570	2	AC021246	Homo sapi
5	250	100.0	96717	9	AL359182	Human DNA
6	250	100.0	175064	2	AC012230	Homo sapi
7	250	100.0	183999	6	AX092589	Sequence
8	248.4	99.4	3231	6	AX351029	Sequence
9	248.4	99.4	149034	9	AF275948	Homo sapi
10	246.8	98.7	1167	9	AF258623	Homo sapi
11	246.8	98.7	201144	9	AF287262	Homo sapi
12	226	90.4	69570	2	AC021246	Homo sapi
13	207.4	83.0	90598	2	AC021345	Homo sapi
14	173	69.2	697	9	AF258627	Homo sapi
15	157.4	63.0	2893	6	AX351031	Sequence
16	132.8	53.1	186889	2	AL807243	Mus muscu
17	132.8	53.1	278572	10	AF287263	Rattus no
18	130.6	52.2	145833	2	AC125837	Sequence
19	98	39.2	7260	6	AX253452	Sequence
20	91	36.4	221	6	AX351032	Sequence
21	89.4	35.8	1556	9	AK024328	Homo sapi
22	89	35.5	1750	9	AK022254	Homo sapi
23	75	30.0	9854	6	AX127831	Sequence
24	75	30.0	9854	6	AX139818	Sequence
25	67	26.8	10442	6	AX060713	Sequence
26	67	26.8	10442	6	AX060892	Sequence
27	67	26.8	10442	9	AF285167	Homo sapi
28	60	24.0	10474	6	AX060719	Sequence
29	60	24.0	10474	6	AX060721	Sequence
30	60	24.0	10474	6	AX060898	Sequence
31	60	24.0	10474	6	AX060900	Sequence
32	39	15.6	207959	2	AC129157	Rattus no
33	37.4	15.0	1599	10	RATMCSR	L27081 Rat melanoc
34	37.4	15.0	75499	9	HSJ955G21	ALI21772 Human DNA
35	37.4	15.0	120166	9	AC091390	Sequence
36	37.4	15.0	125108	9	AF030453	Homo sapi
37	37.4	15.0	150542	2	AC097603	Rattus no
38	36.4	14.6	38	6	AX092823	Sequence
39	36.4	14.6	129025	2	AC105817	Rattus no
40	36.2	14.5	700	6	AX182501	Sequence
41	36.2	14.5	38796	9	AC118464	Homo sapi
42	36.2	14.5	122557	2	AC011429	Homo sapi
43	36.2	14.5	191397	2	AC023861	Homo sapi
44	36.2	14.5	201635	10	AL611931	Mouse DNA
45	36.2	14.5	220087	2	AL662818	Mus muscu

ALIGNMENTS

RESULT 1
HS252201
LOCUS
DEFINITION
HSA252201: Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, 5'UTR and promoter region.
ACCESSION
AJ252201
VERSION
AJ252201.1
KEYWORDS
ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1167)
Porsch-Ozcurumez, M., Langmann, T., Helmerl, S., Borsukova, H.,
Kaminski, W.E., Drobnik, W., Hoyer, C., Schumacher, C. and Schmitz, G.
1167 bp DNA linear PRI 10-APR-2001

TITLE2
The zinc finger protein 202 (ZNF202) is a transcriptional repressor of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene expression and a modulator of cellular lipid efflux

JOURNAL
MEDLINE
21192304
PUBMED
11279031
REFERENCE
2 (bases 1 to 1167)
Porsch-Oezcuernomez, M.K.,
Direct Submission
Submitted (05-JAN-2000) Porsch-Oezcuernomez M.K., Institute for
Clinical Chemistry, University of Regensburg,
F.R.G. Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY

Location/Qualifiers

FEATURES
source
1..1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q22-31"
/cell_type="leukocyte"
1..1167
/gene="ABC-1"
/gene="ABC-1"
1..942
/function="cholesterol efflux regulatory protein"
896..900
/gene="ABC-1"
1148..1167
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
BASE COUNT 278 a 315 c 327 g 247 t
ORIGIN
Query Match 100.0%; Score 250; DB 9; Length 1167;
Best Local Similarity 100.0%; Pred. No. 7e-58; 0; Gaps 0;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS
AX060715
DEFINITION
Sequence 3 from Patent WO0078972.
ACCESSION
AX060715
VERSION
AX060715.1 GI:12406104
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
AUTHORS
Lawn, R.M., Wade, D., and Garvin, M.
TITLE
Regulation with binding cassette transporter protein abc1
JOURNAL
PATENT: WO 0078972-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)

FEATURES
source
1..1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 278 a 315 c 327 g 247 t
ORIGIN
Query Match 100.0%; Score 250; DB 9; Length 1167;
Best Local Similarity 100.0%; Pred. No. 7e-58; 0; Gaps 0;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS
AX060715
DEFINITION
Sequence 3 from Patent WO0078972.
ACCESSION
AX060715
VERSION
AX060715.1 GI:12406104
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
AUTHORS
Lawn, R.M., Wade, D., and Garvin, M.
TITLE
Regulation with binding cassette transporter protein abc1
JOURNAL
PATENT: WO 0078972-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)

FEATURES
source
1..1643
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 370 a 413 c 457 g 403 t
ORIGIN

Query Match 100.0%; Score 250; DB 6; Length 1643;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS
AX060894
DEFINITION
Sequence 3 from Patent WO0078971.
ACCESSION
AX060894
VERSION
AX060894.1 GI:12406271
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
AUTHORS
Lawn, R.M., Wade, D., Oram, J.F., and Garvin, M.
TITLE
ATP binding cassette transporter protein abc1 polypeptides
JOURNAL
PATENT: WO 0078971-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)

FEATURES
source
1..1643
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 370 a 413 c 457 g 403 t
ORIGIN

Query Match 100.0%; Score 250; DB 6; Length 1643;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS
AX060715
DEFINITION
Sequence 3 from Patent WO0078972.
ACCESSION
AX060715
VERSION
AX060715.1 GI:12406104
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
AUTHORS
Lawn, R.M., Wade, D., and Garvin, M.
TITLE
Regulation with binding cassette transporter protein abc1
JOURNAL
PATENT: WO 0078972-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)


```

Db 1574 TGGGCGGACCGCAGCGGACCGACCCCTTCTCTCCGCGGCTCGGCGAGCGCAGG 1633
Qy 241 CGGGGAGGCTC 250
|||||
Db 1634 CGGGGAGGCTC 1643

RESULT 4
AC021246          AC021246          69570 bp      DNA      linear      HTG 13-JUL-2000
LOCUS            Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.
DEFINITION
ACCESSION        AC021246
VERSION          AC021246.2  GI:9119882
KEYWORDS         HTG; HTGS_PHASES;
SOURCE           Homo sapiens
ORGANISM         Homo sapiens
REFERENCE
AUTHORS          Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
                  Anderson, S., Baldwin, J., Barna, N., Becker, R., Beda, F.,
                  Boguslavskiy, L., Boukheir, B., Brown, A., Burkett, G., Castle, A.,
                  Choquel, Y., Cclangelo, M., Collins, S., Collymore, A., Cooke, P.,
                  Dearellano, K., Dewar, K., Domko, W., Doyle, M., Fenestor, J.,
                  Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
                  Gardina, S., Grant, G., Hagos, B., Hearford, A., Horton, L.,
                  Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
                  Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
                  Macdonald, P., Marquis, N., McEwan, P., McCuck, A., McKernan, K.,
                  McPheeters, R., Meltdrim, J., Meneus, L., Morrow, J., Nay-Ori, J.,
                  Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
                  Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
                  Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
                  Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
                  Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
                  Zimmer, A. and Zody, M.
DIRECT SUBMISSION
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: L_N10
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 871: contig of 871 bp in length
* 872 971: gap of 100 bp
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp

```

```

2905 3745: contig of 841 bp in length
3746 3845: gap of 100 bp
3846 4696: contig of 851 bp in length
4697 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 5540: contig of 800 bp in length
5541 6640: gap of 100 bp
6641 7509: contig of 869 bp in length
7510 7609: gap of 100 bp
7610 8479: contig of 870 bp in length
8480 8579: gap of 100 bp
8580 9430: contig of 851 bp in length
9431 9530: gap of 100 bp
9531 10378: contig of 846 bp in length
10377 10478: gap of 100 bp
10477 11322: contig of 846 bp in length
11323 11422: gap of 100 bp
11423 12302: contig of 880 bp in length
12303 12402: gap of 100 bp
12403 13280: contig of 878 bp in length
13281 13380: gap of 100 bp
13381 14241: contig of 861 bp in length
14242 14341: gap of 100 bp
14342 15196: contig of 855 bp in length
15197 15296: gap of 100 bp
15297 16123: contig of 827 bp in length
16124 16223: gap of 100 bp
16224 17072: contig of 849 bp in length
17073 17172: gap of 100 bp
17173 18041: contig of 869 bp in length
18042 18141: gap of 100 bp
18142 19009: contig of 858 bp in length
19010 19109: gap of 100 bp
19110 19966: contig of 837 bp in length
19967 20066: gap of 100 bp
20067 20921: contig of 855 bp in length
20922 21021: gap of 100 bp
21022 21865: contig of 844 bp in length
21865 21965: gap of 100 bp
21966 22832: contig of 867 bp in length
22833 22932: gap of 100 bp
22933 23780: contig of 848 bp in length
23781 23880: gap of 100 bp
23881 24733: contig of 853 bp in length
24734 24833: gap of 100 bp
24834 25670: contig of 837 bp in length
25671 25770: gap of 100 bp
25771 26621: contig of 851 bp in length
26622 26721: gap of 100 bp
26722 27576: contig of 855 bp in length
27577 27676: gap of 100 bp
27677 28532: contig of 856 bp in length
28533 28632: gap of 100 bp
28633 29492: contig of 860 bp in length
29493 29592: gap of 100 bp
29593 30455: contig of 863 bp in length
30456 30555: gap of 100 bp
30556 31410: contig of 855 bp in length
31411 31510: gap of 100 bp
31511 32368: contig of 858 bp in length
32369 32468: gap of 100 bp
32469 33312: contig of 844 bp in length
33313 33412: gap of 100 bp
33413 34268: contig of 856 bp in length
34269 34368: gap of 100 bp
34369 35204: contig of 836 bp in length
35205 35304: gap of 100 bp
35305 36156: contig of 852 bp in length
36157 36256: gap of 100 bp
36257 37128: contig of 872 bp in length
37129 37228: gap of 100 bp
37229 38083: contig of 855 bp in length

```



```

* 45449 45548: gap of 100 bp
* 45549 48116: contig of 2568 bp in length
* 48117 48216: gap of 100 bp
* 48217 52618: contig of 4402 bp in length
* 52619 52718: gap of 100 bp
* 52719 56592: contig of 3874 bp in length
* 56593 56692: gap of 100 bp
* 56693 59635: contig of 2943 bp in length
* 59636 59735: gap of 100 bp
* 59736 63661: contig of 3926 bp in length
* 63662 63781: gap of 100 bp
* 63782 68437: contig of 4676 bp in length
* 68438 68537: gap of 100 bp
* 68538 71458: contig of 2921 bp in length
* 71459 71558: gap of 100 bp
* 71559 76888: contig of 5330 bp in length
* 76889 76988: gap of 100 bp
* 76989 82113: contig of 5125 bp in length
* 82114 82213: gap of 100 bp
* 82214 88220: contig of 6007 bp in length
* 88221 88320: gap of 100 bp
* 88321 93499: contig of 5179 bp in length
* 93500 93599: gap of 100 bp
* 93600 97501: contig of 4302 bp in length
* 97502 98001: gap of 100 bp
* 98002 103016: contig of 5015 bp in length
* 103017 103116: gap of 100 bp
* 103117 109778: contig of 6662 bp in length
* 109779 109278: gap of 100 bp
* 109279 117307: contig of 8029 bp in length
* 117308 117407: gap of 100 bp
* 117408 124079: contig of 6872 bp in length
* 124080 124179: gap of 100 bp
* 124180 131281: contig of 7102 bp in length
* 131282 131381: gap of 100 bp
* 131382 138059: contig of 6672 bp in length
* 138060 138159: gap of 100 bp
* 138160 145431: contig of 7332 bp in length
* 145492 145591: gap of 100 bp
* 145592 157391: contig of 11800 bp in length
* 157392 157491: gap of 100 bp
* 157492 175064: contig of 17573 bp in length.
FEATURES
    source
        1..175064
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="RP11-1M10"
            /clone="RP11-1M10 Human Male BAC"
        1..13003
            /note="assembly_fragment"
            1104..12634
                /note="assembly_fragment"
            2735..4415
                /note="assembly_fragment"
            4516..5785
                /note="assembly_fragment"
            5866..7879
                /note="assembly_fragment"
            7980..9686
                /note="assembly_fragment"
            9787..12253
                /note="assembly_fragment"
            12354..15228
                /note="assembly_fragment"
            15329..17200
                /note="assembly_fragment"
            17301..20131
                /note="assembly_fragment"
            20232..22587
                /note="assembly_fragment"
            22688..25707
                /note="assembly_fragment"
            25808..26194

```

```

misc_feature
    /note="assembly_fragment"
    28285..31338
misc_feature
    /note="assembly_fragment"
    31439..34299
misc_feature
    /note="assembly_fragment"
    34400..38318
misc_feature
    /note="assembly_fragment"
    38419..42835
misc_feature
    /note="assembly_fragment"
    42936..45448
misc_feature
    /note="assembly_fragment"
    45549..48116
misc_feature
    /note="assembly_fragment"
    48217..52618
misc_feature
    /note="assembly_fragment"
    52719..56592
misc_feature
    /note="assembly_fragment"
    56693..59635
misc_feature
    /note="assembly_fragment"
    59736..63661
misc_feature
    /note="assembly_fragment"
    63762..68437
misc_feature
    /note="assembly_fragment"
    68538..71458
misc_feature
    /note="assembly_fragment"
    clone_end:sp6
    vector_side:left
    71559..76888

Query Match      100.0%; Score 250; DB 2; Length 175064;
Best Local Similarity 100.0%; Pred. No. 4.9e-58;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCCCCGGCTCCACGCTGCTTTCTGCTGAGTGAAGTCACTACATAACAGAGCCGGGAA 60
    |||||||
Db 3503 GGCCCCCGGCTCCACGCTGCTTTCTGCTGAGTGAAGTCACTACATAACAGAGCCGGGAA 3562
    |||||||
QY 61 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
    |||||||
Db 3563 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3622
    |||||||
QY 121 GCCGATCTATAAAGAACTAGTCCCGGCAAAACCCGTAATTCGCGAGGAGAGTGAG 180
    |||||||
Db 3623 GCCGATCTATAAAGAACTAGTCCCGGCAAAACCCGTAATTCGCGAGGAGAGTGAG 3682
    |||||||
QY 181 TGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
    |||||||
Db 3683 TGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3742
    |||||||
QY 241 CGGGGAGGCTC 250
    |||||||
Db 3743 CGGGGAGGCTC 3752

RESULT 7
AX092589
LOCUS AX092589 183999 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 1 from Patent WO0115676.
ACCESSION AX092589
VERSION AX092589.1 GI:13444547
KEYWORDS
SOURCE
    human.
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 183999)
AUTHORS
    Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.
TITLE
    Compositions and methods for modulating xdl cholesterol and
    triglyceride levels
JOURNAL
    Patent: WO 0115676-A 1 08-MAR-2001;
    University of British Columbia (CA) ; Xenon Genetics Inc. (CA)
FEATURES
    source
        1..183999

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 49549 a 37944 c 41170 g 54950 t 386 others
ORIGIN
Query Match 100.0%; Score 250; DB 6; Length 183993;
Best Local Similarity 100.0%; Pred. No. 4.9e-58; Indels 0; Gaps 0;
Matches 250; Conservative 0; Mismatches 0;

QY 1 GGGCCCGGCTCCACGCTTCTTCGTGAGTGAACATACATAAACAGAGCGCGGAA 60
DB 28631 GGGCCCGGCTCCACGCTTCTTCGTGAGTGAACATACATAAACAGAGCGCGGAA 28690

QY 61 CGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 28691 CGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 28750

QY 121 GCCGAATCTATAAAGAACTAGTCCGGGAAAACCCCGTAAATGGCGAGGAGGTGAG 180
DB 28751 GCCGAATCTATAAAGAACTAGTCCGGGAAAACCCCGTAAATGGCGAGGAGGTGAG 28810

QY 181 TGGGCGCGGAGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 28811 TGGGCGCGGAGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 28870

QY 241 CGGGAGGCTC 250
DB 28871 CGGGAGGCTC 28880

RESULT 8
AX351029 AX351029 3231 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 1 from Patent WO0183746.
ACCESSION AX351029
VERSION AX351029.1 GI:18616385
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 1 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
Source
1..3231
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 809 a 773 c 876 g 773 t
ORIGIN
Query Match 99.4%; Score 248.4; DB 6; Length 3231;
Best Local Similarity 99.6%; Pred. No. 1.8e-57;
Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCCCGGCTCCACGCTTCTTCGTGAGTGAACATACATAAACAGAGCGCGGAA 60
DB 2735 GGGCCCGGCTCCACGCTTCTTCGTGAGTGAACATACATAAACAGAGCGCGGAA 2794

QY 61 CGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 2795 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854

QY 121 GCCGAATCTATAAAGAACTAGTCCGGGAAAACCCCGTAAATGGCGAGGAGGTGAG 180
DB 2855 GCCGAATCTATAAAGAACTAGTCCGGGAAAACCCCGTAAATGGCGAGGAGGTGAG 2914

QY 181 TGGGCGCGGAGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 2915 TGGGCGCGGAGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2974
```

```
QY 241 CGGGAGGCTC 250
DB 2975 CGGGAGGCTC 2984

RESULT 9
AF275948 AF275948 149034 bp DNA linear PRI 17-JUL-2000
LOCUS
DEFINITION Homo sapiens ABCAL (ABCA1) gene, complete cds.
ACCESSION AF275948
VERSION AF275948.1 GI:9247085
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 149034)
AUTHORS Santamarina-Fojo, S., Peterson, K., Knapper, C., Qiu, Y., Freeman, L.,
Cheng, J.F., Osorio, J., Remaley, A., Yang, X.P., Haudenschield, C.,
Prades, C., Chimini, G., Blackmon, E., Francois, T., Duverger, N.,
Rubin, E.M., Rosier, M., Deneffe, P., Fredrickson, D.S. and Brewer, H.B.
Jr.
TITLE Complete genomic sequence of the human ABCAL gene: analysis of the
human and mouse ATP-binding cassette A promoter.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
MEDLINE 20345099
PUBMED 10884428
REFERENCE 2 (bases 1 to 149034)
AUTHORS Santamarina-Fojo, S., Peterson, K.M., Knapper, C.L., Freeman, J.A.,
Remaley, A.T., Yang, X.-P., Haudenschield, C.C., Blackmon, E.E.,
Francois, T.L. and Brewer, H.B. Jr.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
FEATURES
Source
1..149034
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1454..148034
/gene="ABCA1"
JcIn(1454..1674,25831..25989,40385..40475,45012..45153,
46423..46541,67729..67850,70831..71007,83960..84052,
89010..89250,91962..92101,92433..92549,96758..96955,
97702..97907,98428..98604,100391..100613,102360..102581,
103642..103846,104951..105064,106862..107033,
108023..108154,109450..109602,109805..109943,
110546..110866,112124..112196,113183..113385,
115053..115101,115287..115410,116806..116954,
118604..118728,119964..120062,123094..123283,
124804..124898,126208..126240,127363..127468,
128943..129017,129539..129708,130939..131116,
133114..133229,133341..133485,134526..134649,
135736..135865,136130..136250,137038..137100,
138008..138114,140469..140610,140982..141116,
142060..142163,142646..142738,143397..143640,
144581..148034)
/gene="ABCA1"
/product="ABCA1"
Join(1454..1674,25831..25923)
/gene="ABCA1"
repeat_region 2406..2678
/rpt_family="Alu"
repeat_region 4240..4509
/rpt_family="Alu"
repeat_region 4598..5270
/rpt_family="Alu"
repeat_region 5563..5839
/rpt_family="Alu"
repeat_region 5615..6877
/rpt_family="Alu"
repeat_region 8800..9046
```

```

/rpt_family="Alu"
998..10279
repeat_region
/rpt_family="Alu"
1859..12133
repeat_region
/rpt_family="Alu"
12810..12902
repeat_region
/rpt_family="Alu"
15220..15399
repeat_region
/rpt_family="Alu"
15508..15787
repeat_region
/rpt_family="Alu"
18500..18860
repeat_region
/rpt_family="Alu"
20950..21206
repeat_region
/rpt_family="Alu"
25038..25314
repeat_region
CDS
Join(25924..25989,40385..40478,45012..45153,46423..46541,
67729..67850,70831..71007,83960..84052,89010..89250,
91962..92101,92433..92549,96758..96955,97702..97907,
98428..98504,100391..100613,102350..102581,103642..103846,
104951..105064,106862..107033,108023..108154,
109450..109602,109806..109943,110646..110866,
112124..112196,113183..113385,115053..115101,
115297..115410,116806..116954,118604..118728,
119964..120062,123094..123283,124804..124898,
126208..126240,127363..127468,128943..129017,
129539..129708,130939..131116,133114..133229,
133341..133485,134526..134649,135736..135865,
138130..138250,137038..137100,138008..138114,
140649..140616,140982..141116,142060..142163,
142646..142738,143397..143640,144581..144721)
/gene="ABCA1"
/cocon_start=1
/product="ABCA1"
/protein_id="AAF86276.1"
/db_xref="GI:9247086"
/translation="KACWQLRLLLKLNKLTFRRTQCLLLEAVPLFLILILSVRL
SYPPVQHCHEPNAKPSAGRLPWQGIICNANPCPRYPPTGPAGVGVGNFKSIV
ARLSDARLLLYSKDTSKDMKVLPTLQQLKSSNLKQDFLDVNETSGFLYH
NLSLPKSTVDKLRADYILHKVLOQYQLHLISLNGSKSEEMQLQDQEVSELGPL
REKLAARVLRNMDILPILRLNSTPFSKELAEATKLLHSLGLAEOLSMR
SWSDMRQVFIYNSSSTQIYAVSIVGHPGEGKIKLSLWEDDNNYKALF
GNGREDEATFTYDTPYCNLMLKNLESSRSRIWKALPLWKLKLYLPTDPAT
ROMAPVNTQFELAVFHELEGMEELSPKIWTFMENQEMDLVRMLDSRONCHWE
COLDGLDWTADQIAPLAKHPEDVQSSGVVTRFAETNOAIRTISKREWCYVNLN
KLEPATVYWLKNSMELLDERKFWAGIVFTGTPGSIPLPHVYKIRMDIDNVPRT
NAIKQYWDGPRADPFEDMRVYWGGEFALQVVEQAIIRVLITGTEKKGVYMQOPEY
PCVDDIFLRVSRSPLEFEMLIWYISVAVIIGVIEKARLKEIRIMLGNLSLW
SWFTSSLPLVYAGLLVILKGNLLPYSDPSVVFVLSVFAVVILOCELSLTF
SRANLAACGGIITYPLLYPLVLCVADYVGTFLKIFAXLLSPVAFGCEYFALPE
EOGIGVQNDLPESPVEDGFNLFTSVSMKLPFLYGVMIWYIEAVFPQYQIPRPW
YFPCTKSYWFGESDEKSHPGSNOKRISEICWEEPTHLKLVSIONAKVYSDGMKV
AVDGLALNEYQITSEFLHNGAGKTTMSILTLGFLPTSGTAYILGKDIRSEMSTIR
QNLGVCQHNVLDMTVEHUIFYARLKGSEKVKAEQNALVGLPSSKLKSKT
SOLSGMQKRLSVAFVGGSKVVIDEPTAGVDPYSRGIWELLIKYKQRTILIST
HMDEADVLGDIATISGKLCQGSLEKLNQLGTYLTLWKDQVSSLSGCRSS
HIVDELKEDSVQSSSDAGLSDHSDTITDVSASINLRKHVSEALVEDIGHEL
TVLVPYAAKAGAFVLEHIDRLSDGLGSSYISETLEELFKVASEGSDYATK
DCTLPARRNRAAGFKQSLRPTEDDADPNDSIDPSERTDLLSGMDGGSYAVK
GNLTKQQLVALLKQLLARSRKGFPAQIVLPAVFVCIALVSLVFPFGKPSLE
LQPMWYNEQITFVSNDAPEDTGLELLNALTDKDPTGRCMEGNIPDTPCQAGEEW
TAPAPNTDLFPQNGWNTQNSPACQSSDKIKRMLPVCPCAGLPPQKQNTA
DILQDITGRNISIDLYVYQIIIAKSLKNKINWNEPYFGFSLGVSNTQALPQSOVN
DAKQKHLKLAQSDADFLSLGRLMTGLDTRNNYAVFNNKQWHAISFLNLIK
NAILRNLQGENPSHYGITAFAHPLNLTQQLSEVAXMTISVDVLVYSCVIFAMSEV
PASVVFLEDERVSKAKHLOFISGVFPIYWLSNFVNDMGNVYVPAFLVILIFCQVQ
KSVYSNTNLEADULLLIGWSTIPMPYASFVFKIPSTAYVILTVNCFIGINGSVA
TVLELFTDNKLNINIDILKSVLEIFPHFLQGRGLIDMKVKNQAMADALRFGNRRYS
PLSWDLVRNLFANAGVGVFFLITVLIOVRFTIRPVPNAKLSPINDEDEDVRRRO
RILDGGQNDILELKTITIKYRKRKPAVDRCVIGPPGFCGLLGVNAGKSSYTKM
LTXDFTYVTRGDAFLNLSINLHEVXKRMGYCOPDAILTELLIGREHVEFFALLRGV
PEXEVGVKAGVIRKLGLVIRGKYAGNYSGGNKRKLSTAMALIGLPPVVFLLDEPTTG

```

```

Query Match 99.4% Score 248.4; DB 9; Length 149034;
Best Local Similarity 99.6%; Pred. No. 1.4e-57;
Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGCCCGGCTCCACGTCTTCTTCTGAGTACATACACAGAGCCGGGAA 60
|||||
Db 1295 GGGCCCGGCTCCACGTCTTCTTCTGAGTACATACACAGAGCCGGGAA 1354
|||||

Qy 61 CGGGCGGGGAGGAGGAGCAGAGCTTGACCGATAGTACCTCTGGCTGGTGA 129
|||||

```

```

Db 1355 GGGGGGGGAGGAGGAGGAGGAGGAGGCTTTGACCGATAGTAACTCTCGGCTCGGTGCA 1414
QY 121 GCCGAATCTATAAAGGAAGTACTAGTCCGGCAAAACCCCGTAATTCGGACGAGAGTGAG 180
Db 1415 GCCGAATCTATAAAGGAAGTACTAGTCCGGCAAAACCCCGTAATTCGGACGAGAGTGAG 1474
QY 181 TGGGCGGGGAGCGGAGAGCGGAGCGGAGCGCTTCTCTCCGGGCTCGGCGAGGCGAGGG 240
Db 1475 TGGGCGGGGAGCGGAGAGCGGAGCGGAGCGGAGCGCTTCTCTCCGGGCTCGGCGAGGCGAGGG 1534
QY 241 CGGGGAGCTC 250
Db 1535 CGGGGAGCTC 1544

RESULT 10
LOCUS AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
promoter and exon 1.
ACCESSION AF258623
VERSION AF258623.2 GI:8677405
KEYWORDS
SEGMENT 1 of 4
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Acuizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of hABCL gene 5' end; additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 224 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Acuizerat,B.E., Fielding,C.J. and Kane,J.P.
Direct Submission
Submitted (20-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REFERENCE 3 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Acuizerat,B.E., Fielding,C.J. and Kane,J.P.
Direct Submission
Submitted (23-JUN-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REMARK Sequence update by submitter
COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.
FEATURES
source
1..1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
224..844
/promoter
845..1147
/exon
/number="1"
BASE COUNT 278 a 313 c 328 g 248 t
ORIGIN
Query Match 98.7%; Score 246.8; DB 9; Length 1167;
Best Local Similarity 99.2%; Pred. No. 5.2e-57;
Matches 248; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCGCGGCTCCACGCTGCTTTGCTGAGTGACTGAACACTACATAACAGCGCGGGA 60
Db 768 GGGCGCGGCTCCACGCTGCTTTGCTGAGTGACTGAACACTACATAACAGCGCGGGA 827
QY 61 GGGCGCGGAGGAGGAGGAGGAGGAGGCTTTGACCGATAGTAACTCTCGGCTCGGTGCA 120

```

```

Db 828 GGGGGCGGGAGGAGGAGGAGGAGGAGGCTTTGACCGATAGTAACTCTCGGCTCGGTGCA 887
QY 121 GCCGAATCTATAAAGGAAGTACTAGTCCGGCAAAACCCCGTAATTCGGACGAGAGTGAG 180
Db 888 GCCGAATCTATAAAGGAAGTACTAGTCCGGCAAAACCCCGTAATTCGGACGAGAGTGAG 947
QY 181 TGGGCGGGGAGCGGAGAGCGGAGCGGAGCGCTTCTCTCCGGGCTCGGCGAGGCGAGGG 240
Db 948 TGGGCGGGGAGCGGAGAGCGGAGCGGAGCGGAGCGCTTCTCTCCGGGCTCGGCGAGGCGAGGG 1007
QY 241 CGGGGAGCTC 250
Db 1008 CGGGGAGCTC 1017

RESULT 11
LOCUS AF287262 20144 bp DNA linear PRI 29-APR-2001
DEFINITION Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
and SNAP protein genes, complete cds.
ACCESSION AF287262
VERSION AF287262.1 GI:13876612
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 20144)
AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
TITLE Human and mouse ABCA1 comparative sequencing and transgenesis
studies revealing novel regulatory sequences
JOURNAL Genomics 73 (1), 66-76 (2001)
MEDLINE 21251004
PUBMED 11352367
REFERENCE 2 (bases 1 to 20144)
AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
Direct Submission
Submitted (13-JUL-2000) Genome Science Department, Lawrence
Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
CA 94720, USA
FEATURES
Location/Qualifiers
1..20144
/organism="Homo sapiens"
/db_xref="taxon:9606"
33931..181457
/gene="ABCA1"
join(33931..34151,58319..58476,72893..72986,77520..77661,
78933..79049,101048..101189,104152..104328,117287..117379,
122329..122569,125285..125424,125754..125870,
130079..130276,131023..131228,131749..131925,
133708..133930,135677..135898,136959..137163,
138268..138381,140179..140350,141340..141471,
142777..142919,143123..143260,143963..144183,
145441..145513,146657..146859,148527..148575,
148771..148884,150280..150428,152078..152202,
153438..153536,156568..156757,158278..158372,
159682..159714,160837..160942,162417..162491,
163013..163182,164413..164590,166588..166703,
166815..166959,168505..168628,169715..169844,
170109..170223,171017..171079,171987..172093,
174448..174589,174961..175095,176039..176142,
176635..176717,177376..177619,178560..181457)
/gene="ABCA1"
/product="ATP-binding cassette 1 sub-family A member 1"
33931..34151
/gene="ABCA1"
/number=1
58319..58476
/gene="ABCA1"
/number=2
join(58411..58476,72893..72986,77520..77661,78933..79049,
101048..101189,104152..104328,117287..117379,

```


* NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1	871:	contig of 871 bp in length
872	971:	gap of 100 bp
972	1834:	contig of 883 bp in length
1835	1934:	gap of 100 bp
1935	2804:	contig of 870 bp in length
2805	2904:	gap of 100 bp
2905	3745:	contig of 841 bp in length
3746	3845:	gap of 100 bp
3846	4636:	contig of 851 bp in length
4637	4796:	gap of 100 bp
4797	5640:	contig of 844 bp in length
5641	5740:	gap of 100 bp
5741	6540:	contig of 800 bp in length
6541	6640:	gap of 100 bp
6641	7509:	contig of 869 bp in length
7510	7609:	gap of 100 bp
7610	8475:	contig of 870 bp in length
8480	8579:	gap of 100 bp
8580	9430:	contig of 831 bp in length
9431	9530:	gap of 100 bp
9531	10376:	contig of 845 bp in length
10377	10476:	gap of 100 bp
10477	11322:	contig of 846 bp in length
11323	11422:	gap of 100 bp
11423	12302:	contig of 880 bp in length
12303	12402:	gap of 100 bp
12403	13280:	contig of 878 bp in length
13281	13380:	gap of 100 bp
13381	14241:	contig of 861 bp in length
14242	14341:	gap of 100 bp
14342	15196:	contig of 855 bp in length
15197	15296:	gap of 100 bp
15297	16123:	contig of 827 bp in length
16124	16223:	gap of 100 bp
16224	17072:	contig of 849 bp in length
17073	17172:	gap of 100 bp
17173	18041:	contig of 869 bp in length
18042	18141:	gap of 100 bp
18142	19009:	contig of 868 bp in length
19010	19109:	gap of 100 bp
19110	19968:	contig of 857 bp in length
19967	20066:	gap of 100 bp
20067	20921:	contig of 855 bp in length
20922	21021:	gap of 100 bp
21022	21865:	contig of 844 bp in length
21865	21965:	gap of 100 bp
21965	22832:	contig of 867 bp in length
22833	22932:	gap of 100 bp
22933	23786:	contig of 848 bp in length
23781	23880:	gap of 100 bp
23881	24733:	contig of 853 bp in length
24734	24833:	gap of 100 bp
24834	25670:	contig of 837 bp in length
25671	25770:	gap of 100 bp
25771	26621:	contig of 851 bp in length
26622	26721:	gap of 100 bp
26722	27576:	contig of 855 bp in length
27577	27676:	gap of 100 bp
27677	28532:	contig of 856 bp in length
28533	28632:	gap of 100 bp
28633	29492:	contig of 860 bp in length

```

* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47598: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 881 bp in length
* 55230 55329: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 883 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp

```

```

* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match          90.4%; Score 226; DB 2; Length 69570;
Best Local Similarity 94.0%; Pred. No. 1.9e-51;
Matches 235; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGGCCCGGCTCCAGTSCCTTGTGCTGAGTGAAGTAACTACATATACAGAGCCGGGAA 60
Db 67570 GGGCCCGGCTACAGTCTTTGTGCTGAGTGAAGTAACTACATATACAGAGCCGGCAA 67511
QY 61 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 67510 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 67451
QY 121 GCCGAATCTATAAAGGAAGTACTAGTCCCGGCAAAACCCGTAATTCGGAGCGAGAGTGG 180
Db 67450 GCCGAATCTATAAAGGAAGTACTAGTCCCGGCAAAACCCGTAATTCGGAGCGAGAGTGG 67391
QY 181 TGGGGCGGGACCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 67390 TGGGGCGGGACCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 67331
QY 241 CGGGGAGGCTC 250
Db 67330 CGGGGAGGCTC 67321

RESULT 13
LOCUS AC021345/c
DEFINITION Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
i (bases 1 to 90698)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Balciun,J., Barna,N., Becker,Y.R., Bida,F.,
Boguslavskiy,L., Boukhvalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrel,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardys,S., Grant,G., Hago,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lebeck,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheters,R., Meldrim,J., Menes,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4483
 Center clone name: 24_J_9

* NOTE: This record contains 92 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * over-representation among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1
 * 911 1010: gap of 100 bp in length
 * 1011 1873: contig of 863 bp in length
 * 1874 1973: gap of 100 bp
 * 1974 2824: contig of 851 bp in length
 * 2825 2924: gap of 100 bp in length
 * 2925 3802: contig of 878 bp in length
 * 3803 3902: gap of 100 bp
 * 3903 4816: contig of 914 bp in length
 * 4817 4916: gap of 100 bp
 * 4917 5759: contig of 843 bp in length
 * 5760 5859: gap of 100 bp in length
 * 5860 6764: contig of 905 bp in length
 * 6765 6864: gap of 100 bp
 * 6865 7747: contig of 883 bp in length
 * 7748 7847: gap of 100 bp
 * 7848 8755: contig of 908 bp in length
 * 8756 8855: gap of 100 bp
 * 8856 9753: contig of 898 bp in length
 * 9754 9855: gap of 100 bp
 * 9854 10757: contig of 904 bp in length
 * 10758 10857: gap of 100 bp
 * 10858 11732: contig of 875 bp in length
 * 11733 11832: gap of 100 bp
 * 11833 12739: contig of 907 bp in length
 * 12740 12839: gap of 100 bp
 * 12840 13710: contig of 871 bp in length
 * 13711 13810: gap of 100 bp
 * 13811 14684: contig of 874 bp in length
 * 14685 14784: gap of 100 bp
 * 14785 15662: contig of 878 bp in length
 * 15663 15762: gap of 100 bp
 * 15763 16677: contig of 915 bp in length
 * 16678 16777: gap of 100 bp
 * 16778 17678: contig of 901 bp in length
 * 17679 17778: gap of 100 bp
 * 17779 18679: contig of 901 bp in length
 * 18680 18779: gap of 100 bp
 * 18780 19632: contig of 853 bp in length
 * 19633 19732: gap of 100 bp
 * 19733 20634: contig of 902 bp in length
 * 20635 20734: gap of 100 bp
 * 20735 21620: contig of 886 bp in length
 * 21621 21720: gap of 100 bp
 * 21721 22579: contig of 859 bp in length
 * 22580 22679: gap of 100 bp
 * 22680 23568: contig of 889 bp in length
 * 23569 23668: gap of 100 bp
 * 23669 24554: contig of 886 bp in length
 * 24555 24654: gap of 100 bp
 * 24655 25521: contig of 867 bp in length
 * 25522 25621: gap of 100 bp

* 25622 26487: contig of 866 bp in length
 * 26488 26587: gap of 100 bp
 * 26588 27464: contig of 877 bp in length
 * 27465 27564: gap of 100 bp
 * 27565 28466: contig of 902 bp in length
 * 28467 28566: gap of 100 bp
 * 28567 29464: contig of 898 bp in length
 * 29465 29564: gap of 100 bp
 * 29565 30447: contig of 883 bp in length
 * 30448 30547: gap of 100 bp
 * 30548 31453: contig of 906 bp in length
 * 31454 31553: gap of 100 bp
 * 31554 32452: contig of 899 bp in length
 * 32453 32552: gap of 100 bp
 * 32553 33447: contig of 895 bp in length
 * 33448 33547: gap of 100 bp
 * 33548 34435: contig of 888 bp in length
 * 34436 34535: gap of 100 bp
 * 34536 35433: contig of 898 bp in length
 * 35434 35533: gap of 100 bp
 * 35534 36440: contig of 907 bp in length
 * 36441 36540: gap of 100 bp
 * 36541 37422: contig of 882 bp in length
 * 37423 37522: gap of 100 bp
 * 37523 38402: contig of 880 bp in length
 * 38403 38502: gap of 100 bp
 * 38503 39380: contig of 878 bp in length
 * 39381 39480: gap of 100 bp
 * 39481 40407: contig of 927 bp in length
 * 40408 40507: gap of 100 bp
 * 40508 41405: contig of 898 bp in length
 * 41406 41505: gap of 100 bp
 * 41506 42399: contig of 894 bp in length
 * 42400 42499: gap of 100 bp
 * 42500 43357: contig of 858 bp in length
 * 43358 43457: gap of 100 bp
 * 43458 44356: contig of 899 bp in length
 * 44357 44456: gap of 100 bp
 * 44457 45325: contig of 869 bp in length
 * 45326 45425: gap of 100 bp
 * 45426 46305: contig of 860 bp in length
 * 46306 46405: gap of 100 bp
 * 46406 47302: contig of 897 bp in length
 * 47303 47402: gap of 100 bp
 * 47403 48393: contig of 891 bp in length
 * 48394 48393: gap of 100 bp
 * 48394 49280: contig of 887 bp in length
 * 49281 49380: gap of 100 bp
 * 49381 50250: contig of 870 bp in length
 * 50251 50350: gap of 100 bp
 * 50351 51209: contig of 859 bp in length
 * 51210 51309: gap of 100 bp
 * 51310 52221: contig of 912 bp in length
 * 52222 52321: gap of 100 bp
 * 52322 53205: contig of 884 bp in length
 * 53206 53305: gap of 100 bp
 * 53306 54225: contig of 920 bp in length
 * 54226 54325: gap of 100 bp
 * 54326 55206: contig of 881 bp in length
 * 55207 55306: gap of 100 bp
 * 55307 56185: contig of 859 bp in length
 * 56186 56285: gap of 100 bp
 * 56286 57124: contig of 859 bp in length
 * 57125 57224: gap of 100 bp
 * 57225 58121: contig of 897 bp in length
 * 58122 58221: gap of 100 bp
 * 58222 59075: contig of 854 bp in length
 * 59076 59175: gap of 100 bp
 * 59176 60058: contig of 883 bp in length
 * 60059 60158: gap of 100 bp
 * 60159 61067: contig of 909 bp in length
 * 61068 61167: gap of 100 bp
 * 61168 62051: contig of 884 bp in length

```

* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67547: contig of 893 bp in length
* 67548 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match      83.0%; Score 207.4; DB 2; Length 50698;
Best Local Similarity 94.4%; Pred. No. 2.3e-46;
Matches 236; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 1 GGGCCCGGTCACGCTGTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 6C
Db 87691 GGGCCCGGCTCCACGCTGTTCTGCTGAGTGTGCTGACTACATAACAGAGAGAGAG 87633

QY 61 CGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 87632 CGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 87573

QY 121 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Db 87572 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 87514

QY 181 TGGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 87513 TGGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 87454

QY 241 CGGGGAGCTC 250
Db 87453 CGGGGAGCTC 87444

RESULT 14
AP258627
LOCUS      Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA, partial cds.
DEFINITION
ACCESSION AP258627
VERSION    AP258627.1 GI:7765707
KEYWORDS
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fleiding,C.J. and Kane,J.P.
TITLE      Analysis of hABC1 gene 5' end: additional peptide sequence, promoter region, and four polymorphisms
JOURNAL    Biochem. Biophys. Res. Commun. 271 (2000) in press
REFERENCE  2 (bases 1 to 697)
AUTHORS    Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fleiding,C.J. and Kane,J.P.
TITLE      Submitted (19-APR-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA
JOURNAL
FEATURES   source
            ..697
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="9"

QY 1 GGGCCCGGCTCCACGCTGTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 60
Db 2735 GGGCCCGGCTCCACGCTGTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 2794

QY 61 CGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 2795 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854

QY 121 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 159
Db 2855 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2893

/imap="sc31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABCI"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AAF69513.1"
/db_xref="GI:7769708"
/translation="MACPQLRLILWKNLIFRRQTCLLEVAWPLFIFILILSVRL
SYPPYEQHECHFPNKAMPAGTLPWVGIIICNANPCFRVPTPGAPGVGNFNS"
BASE COUNT 152 a 198 c 190 g 156 t 1 others
ORIGIN
Query Match      69.2%; Score 173; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 8e-37;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 AGAGCAGGCTTTGACCATAGTACCTCTCGCTCGGTCGAGCGGCAATCTATAAAG 137
Db 1 AGAGCAGGCTTTGACCATAGTACCTCTCGCTCGGTCGAGCGGCAATCTATAAAG 60

QY 138 AACTAGTCCCGCAAAACCCGCTTAATTCGAGCGAGAGTGAAGTGGGCGCGGACCCGCA 197
Db 61 AACTAGTCCCGCAAAACCCGCTTAATTCGAGCGAGAGTGAAGTGGGCGCGGACCCGCA 120

QY 198 GAGCGAGCGACCCCTTCTCTCCGGGCTCGCGCAGGCGGCGGAGGCTC 250
Db 121 GAGCGAGCGACCCCTTCTCTCCGGGCTCGCGCAGGCGGCGGAGGCTC 173

RESULT 15
AX351031
LOCUS      Sequence 3 from Patent WO183746.
DEFINITION
ACCESSION AX351031
VERSION    AX351031.1 GI:18616387
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P., Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.
TITLE      Regulatory nucleic acid sequences of the abci gene
JOURNAL    Patent: WO 0183746-A 3 08-NOV-2001;
            Aventis Pharma S.A. (FR)
FEATURES   Location/Qualifiers
            source
            1..2893
            /organism="Homo sapiens"
            /db_xref="taxon:9606"

BASE COUNT 745 a 581 c 769 g 698 t
ORIGIN
Query Match      63.0%; Score 157.4; DB 6; Length 2893;
Best Local Similarity 99.4%; Pred. No. 1.3e-32;
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGCCCGGCTCCACGCTGTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 60
Db 2735 GGGCCCGGCTCCACGCTGTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 2794

QY 61 CGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 2795 GGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854

QY 121 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 159
Db 2855 GCCGATCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2893

```

Search completed: February 4, 2003, 01:37:29
Job time : 1021.76 secs

GenCore version 5.1.3
Copyr:ght (c) 1993 - 2003 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:46:48 ; Search time 68.2784 Seconds
(without alignments)
8245.650 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1394_1643

Perfect score: 250

Sequence: 1 gggccggcgtccacgtgct.....caggccaggggcgggagctc 250

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	100.0	1197	24	ABL56400 Human large ATP-bi
2	250	100.0	1643	22	AAF24681 Nucleotide sequenc
3	250	100.0	1643	22	AAF24703 Nucleotide sequenc
4	250	100.0	18399	22	AAF92831 Human ABL1 genomic
5	248.4	99.4	3231	24	AAD37265 Human ABL1 transcr
6	174.4	69.8	2910	21	AAF37267 Human ABL1 transcr
7	119.4	47.8	227	21	AAC05615 Human secreted pro
8	98	39.2	7260	22	AAD21326 Human ATP binding
9	98	39.2	7260	22	AAI70315 Human ATP binding

10	97	36.4	221	24	AAD37268 Human ABL1 gene ex
11	89.4	35.8	736	22	AAH07432 Human cDNA clone (
12	89.4	35.8	1556	22	AAH18606 Human cDNA sequenc
13	89	35.6	1750	22	AAH04729 Human cDNA clone (
14	89	35.6	1750	22	AAH17451 Human cDNA sequenc
15	77.2	30.9	7281	22	AAK51683 Human polynucleoti
16	75.4	30.2	7086	22	ABA09200 Human ABL1 homolo
17	75.4	30.2	7086	22	AAK52667 Human polynucleoti
18	75	30.0	9854	22	AAK06121 Human ABL1 DNA seq
19	67	26.8	10442	22	AAF24680 Nucleotide sequenc
20	67	26.8	10442	22	AAF24702 Nucleotide sequenc
21	60	24.0	10474	22	AAF24585 Nucleotide sequenc
22	60	24.0	10474	22	AAF24686 Nucleotide sequenc
23	60	24.0	10474	22	AAF24707 Nucleotide sequenc
24	60	24.0	10474	22	AAF24708 Rat melanocortin r
25	37.4	15.0	1599	16	AAO97701 ABL1 polymorphism
26	36.4	14.6	38	22	AAF93064 Human inflammatory
27	36.2	14.5	700	22	AAH92242 DNA encoding novel
28	34.4	13.8	669	23	AAK83553 Human cDNA 5'-end
29	34	13.6	524	22	AAK92108 Human cDNA clone r
30	34	13.6	524	22	AAK93547 Human full-length
31	34	13.6	1769	22	AAK94827 Human cDNA clone (
32	33.2	13.3	595	22	AAH04659 Human extracellular
33	33.2	13.3	1235	22	AAH08064 Nucleotide sequenc
34	33.2	13.3	1877	22	AAH77797 Human stomach canc
35	33.2	13.3	1877	22	AAI93858 Human cDNA sequenc
36	33.2	13.3	1877	22	AAH15566 Human agarecan deg
37	32.6	13.0	3250	20	AAK00726 Human ORFX ORF90
38	32.6	13.0	5530	21	AAK75435 Aspergillus oryzae
39	31.8	12.7	1048	21	AAF14080 Human protein HP03
40	31.6	12.6	1260	22	AAF28692 Human cancer agent
41	31.4	12.6	1704	22	AAK60816 Melanocortin-5 rec
42	31.2	12.5	975	18	AAK68791 Melanocortin-5 rec
43	31.2	12.5	975	19	AAK62353 Mouse melanocortin
44	31.2	12.5	975	19	AAV03916 Mouse melanocortin
45	31.2	12.5	975	19	AAV06405

ALIGNMENTS

RESULT 1
ABL58400
ID ABL58400 standard; DNA; 1197 BP.

NC ABL58400;

XX 30-JUL-2002 (first entry)

DI Human large ATP-binding cassette transporter 1(hABL1) promoter sequence.

DE Human large ATP-binding cassette transporter 1; ABL1; promoter;

KW Human; large ATP-binding cassette transporter 1; ABL1; promoter;

KW antiarteriosclerotic; gene transfer; transactivator; ds.

XX Homo sapiens.

OS Homo sapiens.

PN WO200183506-A1.

PD 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13654.

PR 28-APR-2000; 2000US-0560372.

XX (UYCO) UNIV COLUMBIA NEW YORK.

PI Tail AP;

DR WPI; 2002-049334/06.

XX Novel isolated human large ATP-binding cassette transporter 1 promoter
PI capable of directing transcription of heterologous coding sequence
PT positioned downstream to it, useful for expressing foreign DNA in host

```
PT cells -
XX
PS Claim 1; Fig 3; 68pp; English.
XX
CC The invention relates to an isolated human large ATP-binding cassette
CC transporter 1 (ABCI) promoter capable of directing transcription of
CC heterologous coding sequence positioned downstream to it. The ABCI
CC promoter is useful for expressing foreign DNA in a host cell, by
CC introducing into the host cell a gene transfer vector comprising the
CC promoter operably linked to a foreign DNA encoding a desired polypeptide
CC or RNA, where the foreign DNA is expressed. The gene transfer can be
CC introduced into the host cell by adenovirus infection, liposome-mediated
CC transfer, topical application to the cell or microinjection. The gene
CC transfer vector encodes and expresses a reporter molecule. The method
CC further involves introducing into the cell a gene transfer vector
CC comprising a nucleic acid segment encoding a transactivator protein
CC capable of upregulating the ABCI promoter, or contacting the cell with
CC the transactivator protein, or an agonist of the transactivator protein.
CC Modulators of human ABCI gene expression are useful for treating
CC atherosclerosis. The present sequence represents the ABCI promoter.
XX
SQ Sequence 1197 BP; 284 A; 314 C; 328 G; 271 T; 0 other;
Query Match 100.0%; Score 250; DB 24; Length 1197;
Best Local Similarity 100.0%; Pred. No. 1.7e-66;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCACGTGCTTCTCTGAGTGACTGACTACATAACACAGAGCGCGGAA 60
DB 932 GGGCCCGGGTCCACGTGCTTCTCTGAGTGACTGACTACATAACACAGAGCGCGGAA 991
QY 61 CGGGCGGGGAGGAGACAGCGCTTTGACCGATAGTACCTCTGCGCTCGGTGCA 120
DB 992 CGGGCGGGGAGGAGGAGACAGCGCTTTGACCGATAGTACCTCTGCGCTCGGTGCA 1051
QY 121 GCCGAATCTATAAAGGAACTAGTCCCGGCAAAACCCGTAATTCGAGGCGAGAGTGAG 180
DB 1052 GCCGANTCTATAAAGGAACTAGTCCCGGCAAAACCCGTAATTCGAGGCGAGAGTGAG 1111
QY 181 TGGGCGCGGGACCCGACGAGCGGACCGACCTTCTCTCCCGGGTGGCGGAGGCGAGG 240
DB 1112 TGGGCGCGGGACCCGACGAGCGGACCGACCTTCTCTCCCGGGTGGCGGAGGCGAGG 1171
QY 241 CGGGGAGGCTC 250
DB 1172 CGGGGAGGCTC 1181
RESULT 2
AAF24681
ID AAF24681 standard; DNA; 1643 BP.
XX
AC AAF24681;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the 5' flanking region of the human ABCI gene.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABCI;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
PN WO200078972-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16765.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
```

```
PR 19-NOV-1999; 99US-0166573.
XX (CVTH-) CV THERAPEUTICS INC.
XX
PI Lawn RM, Wade D, Garvin M;
XX WPI; 2001-137812/14.
XX
PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PT useful for the development of agents for the treatment of heart disease
PT and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
XX Claim 1; Page 143-144; 2:5pp; English.
XX
CC The present sequence represents the 5' flanking region of the human
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABCI
CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABCI is a pivotal
CC protein in the apolipoprotein-mediated mobilisation of intracellular
CC cholesterol stores. ABCI is defective in Tangier disease, a genetic
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABCI
CC gene is localised to chromosome 9q22-9q31. The ABCI genes and proteins
CC are useful for developing pharmaceutical agents for the treatment of
CC heart disease and other disorders associated with hypercholesterolemia
CC and atherosclerosis. The genes are useful for developing screening assays
CC to screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;
Query Match 100.0%; Score 250; DB 22; Length 1643;
Best Local Similarity 100.0%; Pred. No. 1.8e-66;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCACGTGCTTCTCTGAGTGACTGACTACATAACACAGAGCGCGGAA 60
DB 1394 GGGCCCGGGTCCACGTGCTTCTCTGAGTGACTGACTACATAACACAGAGCGCGGAA 1453
QY 61 CGGGCGGGGAGGAGGAGACAGCGCTTTGACCGATAGTACCTCTGCGCTCGGTGCA 120
DB 1454 CGGGCGGGGAGGAGGAGACAGCGCTTTGACCGATAGTACCTCTGCGCTCGGTGCA 1513
QY 121 GCCGAATCTATAAAGGAACTAGTCCCGGCAAAACCCGTAATTCGAGGCGAGAGTGAG 180
DB 1514 GCCGAATCTATAAAGGAACTAGTCCCGGCAAAACCCGTAATTCGAGGCGAGAGTGAG 1573
QY 181 TGGGCGCGGGACCCGACGAGCGGACCGACCTTCTCTCCCGGGTGGCGGAGGCGAGG 240
DB 1574 TGGGCGCGGGACCCGACGAGCGGACCGACCTTCTCTCCCGGGTGGCGGAGGCGAGG 1633
QY 241 CGGGGAGGCTC 250
DB 1634 CGGGGAGGCTC 1643
RESULT 3
AAF24703
ID AAF24703 standard; DNA; 1643 BP.
XX
AC AAF24703;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the 5' flanking region of the human ABCI gene.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABCI;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
```



```
AAD37265
ID AAD37265 standard; DNA: 3231 BP.
XX
AC AAD37265;
XX
XX 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #1.
XX
XX Human: ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200183746-A2.
XX
XX 08-NOV-2001.
XX
XX 02-MAY-2001: 2001-WO-EP05488.
XX
XX 02-MAY-2000: 2000US-201280P.
XX
XX (AVET ) AVENTIS PHARMA SA.
XX
XX Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
XX Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
XX WPI: 2002-154404/20.
XX
XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
XX (ABC1) and screening for candidate modulatory compounds or substances
XX
XX Claim 1: Page 130-131; 152pp; English.
XX
XX The invention relates a nucleic acid which is capable of regulating the
XX transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
XX casual gene for pathologies linked to a dysfunctioning of cholesterol
XX metabolism, including diseases such as atherosclerosis. Polynucleotides
XX of the invention are used to screen candidate molecules or substances
XX that are capable of modulating the transcription of the ABC1 gene. They
XX are used in antisense therapy. Compositions comprising sequences of the
XX invention are used to treat hypercholesterolaemia and atherosclerosis.
XX The present sequence is human ABC1 transcription regulating DNA.
XX
XX Sequence 3231 Bp; 809 A; 773 C; 876 G; 773 T; 0 other;
XX
Query Match 99.4%; Score 248.4; DB 24; Length 3231;
Best Local Similarity 99.6%; Pred. No. 7e-66;
Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCGGGTCCACGTGCTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 60
Db |||||
2735 GGGCCCGGGTCCACGTGCTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAA 2794
QY 61 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db |||||
2795 GGGCGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854
QY 121 GCCCAATCTATATAAAGGAAGTCTGCTCCGGCAAAACCCCGTAATTGGGAGCGAGGTGAG 180
Db |||||
2855 GCCCAATCTATATAAAGGAAGTCTGCTCCGGCAAAACCCCGTAATTGGGAGCGAGGTGAG 2914
QY 181 TGGGGCGGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 240
Db |||||
2915 TGGGGCGGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2974
QY 241 CGGGGAGCTC 250
Db |||||
2975 CGGGGAGCTC 2984
RESULT 6
```

```
AAD37257
ID AAD37257 standard; DNA: 2910 BP.
XX
AC AAD37257;
XX
XX 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #3.
XX
XX Human: ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX protein_bind 12..23
XX /tag= a
XX /bound_moiety= "LMO2COM/MYOD"
XX protein_bind 97..107
XX /tag= b
XX /bound_moiety= "DeltaEFl"
XX protein_bind 110..125
XX /tag= c
XX /bound_moiety= "S8/NKX2.5"
XX protein_bind 196..211
XX /tag= d
XX /bound_moiety= "S8"
XX protein_bind 228..237
XX /tag= e
XX /bound_moiety= "GATA"
XX protein_bind 399..410
XX /tag= f
XX /bound_moiety= "IK2"
XX protein_bind 412..420
XX /tag= g
XX /bound_moiety= "LYF1"
XX protein_bind 528..539
XX /tag= h
XX /bound_moiety= "LMO2COM/MYOD/DeltaEFl"
XX protein_bind 549..556
XX /tag= i
XX /bound_moiety= "LYF1"
XX protein_bind 558..568
XX /tag= j
XX /bound_moiety= "DeltaEFl"
XX protein_bind 590..596
XX /tag= k
XX /bound_moiety= "LNX2.5"
XX protein_bind 608..620
XX /tag= l
XX /bound_moiety= "NFY/CAAT"
XX protein_bind 708..715
XX /tag= m
XX /bound_moiety= "MZFl"
XX protein_bind 723..730
XX /tag= n
XX /bound_moiety= "MZFl"
XX protein_bind 771..785
XX /tag= o
XX /bound_moiety= "HFH2/SRY/EV11"
XX protein_bind 803..812
XX /tag= p
XX /bound_moiety= "CREBP1/VBP"
XX protein_bind 831..837
XX /tag= r
XX /bound_moiety= "NKX2.5"
XX protein_bind 1076..1089
XX /tag= s
XX /bound_moiety= "GATA"
XX protein_bind 1173..1188
XX /tag= t
XX /bound_moiety= "LXRalpha/DeltaEFl"
XX protein_bind 1189..1209
```

```

FT FT      /tag= u
FT FT      /bound_moiety= "DeltaE1/UYF1/IX2"
FT FT      1483..1491
FT FT      /tag= v
FT FT      /bound_moiety= "AP4"
FT FT      1498..1514
FT FT      /tag= w
FT FT      /bound_moiety= "LMO2-COM/MYOB/deltaE1/E47"
FT FT      1524..1545
FT FT      /tag= x
FT FT      /bound_moiety= "ZID/deltaE1"
FT FT      1597..1607
FT FT      /tag= y
FT FT      /bound_moiety= "DeltaE1"
FT FT      1622..1627
FT FT      /tag= z
FT FT      /bound_moiety= "PPAR"
FT FT      1632..1637
FT FT      /tag= aa
FT FT      /bound_moiety= "PPAR"
FT FT      1685..1698
FT FT      /tag= ab
FT FT      /bound_moiety= "USE/NMYC/MYCMAX"
FT FT      1787..1797
FT FT      /tag= ac
FT FT      /bound_moiety= "DeltaE1"
FT FT      1809..1819
FT FT      /tag= ad
FT FT      /bound_moiety= "DeltaE1"
FT FT      1822..1833
FT FT      /tag= ae
FT FT      /bound_moiety= "SRY"
FT FT      1840..1850
FT FT      /tag= af
FT FT      /bound_moiety= "AP1"
FT FT      1942..1956
FT FT      /tag= ag
FT FT      /bound_moiety= "HNF3beta"
FT FT      1978..1985
FT FT      /tag= ah
FT FT      /bound_moiety= "NKX2.5"
FT FT      2008..2016
FT FT      /tag= ai
FT FT      /bound_moiety= "PPAR/NKX2.5/PPAR"
FT FT      2019..2024
FT FT      /tag= aj
FT FT      /bound_moiety= "PPAR/NKX2.5/PPAR"
FT FT      2051..2059
FT FT      /tag= ak
FT FT      /bound_moiety= "GATA"
FT FT      2104..2111
FT FT      /tag= al
FT FT      /bound_moiety= "SOX5"
FT FT      2114..2152
FT FT      /tag= am
FT FT      /bound_moiety= "SRY/HFH/HNF3beta"
FT FT      2221..2228
FT FT      /tag= an
FT FT      /bound_moiety= "MZF1"
FT FT      2234..2249
FT FT      /tag= ao
FT FT      /bound_moiety= "IK2/NFkappaB/CREL"
FT FT      2259..2272
FT FT      /tag= ap
FT FT      /bound_moiety= "LMO2COM/GATA"
FT FT      2289..2306
FT FT      /tag= aq
FT FT      /bound_moiety= "MZF1/SRY"
FT FT      2313..2318
FT FT      /tag= ar
FT FT      /bound_moiety= "PPAR"
FT FT      2321..2326
FT FT      /tag= as

```

```

FT FT      /bound_moiety= "PPAR"
FT FT      2335..2342
FT FT      /tag= at
FT FT      /bound_moiety= "MZF1"
FT FT      2361..2384
FT FT      /tag= au
FT FT      /bound_moiety= "HNF3beta/SRY/EV11"
FT FT      2426..2433
FT FT      /tag= av
FT FT      /bound_moiety= "MZF1"
FT FT      2442..2451
FT FT      /tag= aw
FT FT      /bound_moiety= "AP4"
FT FT      2455..2466
FT FT      /tag= ax
FT FT      /bound_moiety= "SRY"
FT FT      2491..2498
FT FT      /tag= ay
FT FT      /bound_moiety= "STAT"
FT FT      2524..2534
FT FT      /tag= az
FT FT      /bound_moiety= "STAT/PPAR"
FT FT      2536..2541
FT FT      /tag= ba
FT FT      /bound_moiety= "PPAR"
FT FT      2589..2600
FT FT      /tag= bb
FT FT      /bound_moiety= "AP2"
FT FT      2610..2617
FT FT      /tag= bc
FT FT      /bound_moiety= "MZF1"
FT FT      2634..2648
FT FT      /tag= bd
FT FT      /bound_moiety= "LMO2COM/MYOB/E47"
FT FT      2657..2672
FT FT      /tag= be
FT FT      /bound_moiety= "RREB1"
FT FT      2680..2696
FT FT      /tag= bf
FT FT      /bound_moiety= "MZF1/CMYB"
FT FT      2728..2740
FT FT      /tag= bc
FT FT      /bound_moiety= "SP1/GC"
FT FT      2745..2757
FT FT      /tag= bh
FT FT      /bound_moiety= "USE/NMYC/ARNT"
FT FT      2758..2773
FT FT      /tag= bi
FT FT      /bound_moiety= "NFE2A1"
FT FT      2774..2787
FT FT      /tag= bj
FT FT      /bound_moiety= "XFD1/HFH"
FT FT      2794..2806
FT FT      /tag= bk
FT FT      /bound_moiety= "GC/SP1/MZF1"

```

Query Match 69.8%; Score 174.4; DB 24; Length 2910;

Best Local Similarity 99.4%; Pred. No. 2.7e-43;

Matches 175; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GGGCCCCGGCTCCACGTCTTCTGCTGAGTGAAGTAACTAACACAGAGCCCGGAA 60
    |||
Db 2735 GGGCCCCGGCTCCACGTCTTCTGCTGAGTGAAGTAACTAACACAGAGCCCGGAA 2794
    |||

QY 61 GGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
    |||
Db 2795 GGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2854
    |||

QY 121 GCGGAATCTATAAAGGAAGTATGTCGGGCAAAACCCGCTAATTCGAGCGAGAG 176
    |||
Db 2855 GCGGAATCTATAAAGGAAGTATGTCGGGCAAAACCCGCTAATTCGAGCGAGAG 2910
    |||

```

RESULT 7
AAC09615
ID AAC09615 standard; cDNA; 227 BP.
XX AC AAC09615;
XX DT 06-OCT-2000 (first entry)
XX DE Human; secreted protein 5' EST, SEQ ID NO: 13690.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN BP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GIST) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX PS New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 1; SEQ ID 13690; 71bp = CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SQ Sequence 227 BP; 44 A; 65 C; 73 G; 45 T; 0 other;
Query Match 47.8%; Score 119.4; DB 21; Length 227;
Best Local Similarity 99.2%; Pred. No. 7.7e-27;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 130 ATAAAGGAAGTACTCCCGGCAAAACCCCGTAATTCGAGCGAGAGTGGGGCGG 189
Db 1 ATAAAGGAAGTACTCTCGGCAAAACCCCGTAATTCGAGCGAGAGTGGGGCGG 60
QY 190 GACCCGAGAGCGGCGGACCGCTTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCT 249
Db 61 GACCCGAGAGCGGCGGACCGCTTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCT 120
QY 250 C 250
Db 121 C 121

RESULT 8
AAD21326
ID AAD21326 standard; DNA; 7260 BP.
XX

AC AAD21326;
XX 28-JAN-2002 (first entry)
XX Human ATP binding cassette transporter 1 (ABCI) gene.
XX KW Human; ATP binding cassette transporter 1; ABC1; coronary heart disease;
XX KW dermatological; atherosclerosis; cardiovascular; inflammatory disease;
XX KW psoriasis; lipid disorder; antibacterial; septic shock; gene therapy;
XX KW immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT CDS 321..7106
XX FT /*tag= 2
XX FT /product= "human ABC1 protein"
XX PN EP1136552-A1.
XX PD 26-SEP-2001.
XX PF 20-MAR-2000; 2000EP-0105820.
XX PR 20-MAR-2000; 2000EP-0105820.
XX PA (FARB) BAYER AG.
XX PI Schmitz G, Bodzioch M;
XX DR WPI; 2001-640388/74.
XX DR P-PSDB; AAB13022.
XX PT New adenosine triphosphate binding cassette transporter-1 gene
XX PT polymorphisms, useful for diagnosing and treating lipid disorders,
XX PT cardiovascular diseases and inflammatory diseases
XX PS Example 1; Fig 1; 48pp; English.
XX CC The invention relates to four common polymorphisms in the gene encoding
CC ATP-binding cassette transporter-1 (ABCI). ABC1 is associated with
CC decreased ApoA-1 mediated efflux of cholesterol. The polymorphisms in
CC ABC1 directly affects cellular lipid homeostasis, which is a key factor
CC in the atherogenic processes. The ABC1 polymorphisms are useful for
CC diagnosing and treating lipid disorders, cardiovascular diseases
CC (coronary heart disease, atherosclerosis) and inflammatory diseases
CC (psoriasis, lupus erythematosus). The identification of ABC1 as a
CC transporter for interleukin-beta (IL-1beta) identifies this gene as
CC a candidate for treatment of inflammatory diseases including rheumatoid
CC arthritis and septic shock. The present sequence is human ABC1 gene.
XX SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
Query Match 39.2%; Score 98; DB 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 7.4e-20;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 153 AACCCCGTAATTCGAGCGAGAGTGGGGCGGCGGCGGCGGCGGCGGCGGCGG 212
Db 1 AACCCCGTAATTCGAGCGAGAGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 213 TTCTCTCCCGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 250
Db 61 TTCTCTCCCGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 98
RESULT 9
AAI70315
ID AAI70315 standard; cDNA; 7260 BP.
XX AC AAI70315;
XX DT 07-JAN-2002 (first entry)

XX Human ATP binding cassette transporter 1 (ABCI) cDNA.
XX
KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;
KW cholesterol; cardiovascular disease; inflammatory disease;
KW antinflammatory; antipapenic; antipsoriatic; dermatological;
KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
KW polymorphism; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 321..7106
FT FT /*tag= a
FT CDS 501..7106
FT FT /*tag= b
FT FT /*note= "alternative open reading frame of AAI70314"
FT FT replace(976,A)
FT FT /*tag= c
FT FT variation replace(1516,C)
FT FT /*tag= d
FT FT variation replace(2969,G)
FT FT /*tag= e
FT FT variation replace(3836,C)
FT FT /*tag= f
XX
PN EP1136554-A1.
XX
XX 26-SEP-2001.
XX
XX 24-MAR-2000; 2000EP-0106401.
XX
XX 24-MAR-2000; 2000EP-0106401.
XX (FARB) BAYER AG.
XX
XX Schmitz G, Bodzioch M;
XX WPI; 2001-640389/74.
XX P-PSDB; AAM50228.
XX
XX New adenosine triphosphate binding cassette transporter gene
XX polymorphisms, useful for diagnosing and treating lipid disorders,
XX cardiovascular diseases and inflammatory diseases
XX
XX Disclosure; Page 26-28; 41pp; English.
XX
XX The present sequence is that of cDNA encoding the human adenosine
XX triphosphate (ATP) binding cassette transporter 1 (ABCI) protein
XX (see AAM50227). The sequence includes an extended open reading
XX frame (ORF) to that provided by the sequence in AAI70314, using
XX an alternative ATG codon as initiation codon and thereby adding an
XX extra 40 N-terminal amino acids to the encoded ABC1 protein (see
XX AAM50228). The invention provides 4 common polymorphisms in the
XX ABC1 gene. These were identified by sequencing the ABC1 gene in
XX different Tangier kindreds. In the variant genes (numbering as in
XX AAI70314), G is changed to A at position 596, T is changed to C at
XX position 1136, A is changed to G at position 2569 or G is changed
XX to C at position 3456, or any combination of these. All of these
XX polymorphisms alter the amino acid sequence of ABC1 and therefore
XX may affect its function. The 2 most common polymorphisms (G596A
XX and A2569G) are both associated with a decreased in vitro ApoA-I
XX mediated efflux of cholesterol from mononuclear phagocytes, a
XX feature typical of Tangier disease. 3 Of the variants (G596A,
XX A2569G and G3456C) are significantly increased in a population of
XX men having low high density lipoprotein-cholesterol levels and
XX established coronary heart disease (CHD) relative to CHD-free
XX control subjects. The use of the provided ABC1 polymorphisms for
XX the diagnosis and treatment of lipid disorders, cardiovascular
XX diseases, and inflammatory diseases (e.g. psoriasis, lupus
XX erythematoses) is claimed. Modulation of ABC1 transcripts or
XX proteins by antisense or ribozyme technology or RNA decoys is also
XX claimed.

XX
SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
Query Match 39.2%; Score 98; DB 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 7.4e-20;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 153 AAACCCCGTAATTCGGAGCGAGAGTGTGGGCGGCGGAGCCGAGCGAGCCGACCC 212
DB 1 AAACCCCGTAATTCGGAGCGAGAGTGTGGGCGGCGGAGCCGAGCGAGCCGACCC 60
QY 213 TTCTCTCCCGGCTGGCGAGGCGAGGCGGGGAGCTC 250
DB 61 TTCTCTCCCGGCTGGCGAGGCGAGGCGGGGAGCTC 98
RESULT 10
AAD37268
ID AAD37268 standard; DNA; 221 BP.
XX
AC AAD37268;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human ABC1 gene exon 1A.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy;
KW exon 1A; GS.
XX
OS Homo sapiens.
XX
PN WC200183746-A2.
XX
XX 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-EP05488.
XX
XX 02-MAY-2000; 2000US-201280P.
XX
XX (AVET) AVENTIS PHARMA SA.
XX
XX Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
XX Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
XX WPI; 2002-154404/20.
XX
XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
XX (ABCI) and screening for candidate modulatory compounds or substances
XX
XX Claim 4; Page 132; 152pp; English.
XX
XX The invention relates a nucleic acid which is capable of regulating the
XX transcription of human ATP-binding cassette 1 (ABCI) gene, which is a
XX casual gene for pathologies linked to a dysfunctioning of cholesterol
XX metabolism, including diseases such as atherosclerosis. Polynucleotides
XX of the invention are used to screen candidate molecules or substances
XX that are capable of modulating the transcription of the ABC1 gene. They
XX are used in antisense therapy. Compositions comprising sequences of the
XX invention are used to treat hypercholesterolaemia and atherosclerosis.
XX The present sequence is human ABC1 gene exon 1A.
XX
XX Sequence 221 BP; 44 A; 52 C; 73 G; 42 T; 0 other;
Query Match 35.4%; Score 91; DB 24; Length 221;
Best Local Similarity 100.0%; Pred. No. 3.6e-18;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 160 GTAATTGCCAGGCGAGAGTGTGGGCGGCGGAGCCGAGCGAGCCGACCCCTTCTC 219
DB 1 GTAATTGCCAGGCGAGAGTGTGGGCGGCGGAGCCGAGCGAGCCGACCCCTTCTC 50

CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other:

Query Match 35.6%; Score 89; DB 22; Length 1750;
 Best Local Similarity 100.0%; Pred. No. 2.7e-17;
 Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AATTGGCAGCGAGTGAGTGGGGCGGGACCGCGAGAGCGGACCCCTTCTCTCC 221
 DB 1 AATTGGCAGCGAGTGAGTGGGGCGGGACCGCGAGAGCGGACCCCTTCTCTCC 60

QY 222 GGCTCTCGCGAGCGGCGAGGGGGAGCTC 250
 DB 61 GGCTCTCGCGAGCGGCGAGGGGGAGCTC 89

RESULT 15
 AAK51683
 ID AAK51683 standard; cDNA; 7281 BP.
 XX AAK51683;
 AC AAK51683;
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 228.
 XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-JS04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA {HYSE-} HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM78550.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PT
 PS Claim 1: Page 1086-1096; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX
 SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other:

Query Match 30.9%; Score 77.2; DB 22; Length 7281;
 Best Local Similarity 86.7%; Pred. No. 1.7e-13;
 Matches 85; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 153 AAACCCCGTAATTGCGAGCGAGAGTGAGTGGGCGGGACCGCGAGAGCGGACCC 212
 DB 22 AATTCGCGGGTCGACGATTTCGTGAGTGGGCGGGACCGCGAGAGCGGACCC 81

QY 213 TTCTCTCCCGGCTCGCGAGCGGCGAGGGGGAGCTC 250
 DB 82 TTCTCTCCCGGCTCGCGAGCGGCGAGGGGGAGCTC 119

Search completed: February 3, 2003, 16:34:54
 Job time : 165.421 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:39:45 ; Search time 473.043 Seconds
(without alignments)
8559.210 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1394_1643

Perfect score: 250

Sequence: 1 gggcccggtccacgtgt.....cagggcaggcgaggacctc 250

Scoring table: IDENTITY.NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpi:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89.4	35.8	736	9 AUI35588	AUI35588 AUI35588
2	89	35.6	547	9 AL698654	AL698654 DKT2p886N
3	89	35.6	763	9 AUI21731	AUI21731 AUI21731
4	75	30.0	292	14 Z44377	Z44377 #SC:2B081 n
5	75	30.0	1004	11 BC034824	BC034824 Homo sapi
6	60.2	24.1	998	12 BG678861	BG678861 602624760

C	7	39.6	15.8	912	17	CNS006N3	AL0665775 Drosophil
C	8	37.8	15.1	1863	17	GGA200110	AJ233825 Gallus ga
C	9	37.2	14.9	1201	17	CNS016BR	AL106545 Drosophil
C	10	36.8	14.7	1101	17	CNS012JV	AL101653 Drosophil
C	11	36	14.4	1101	17	CNS00GP3	AL073017 Drosophil
C	12	35.4	14.2	925	17	CNS0091P	AL053013 Drosophil
C	13	35	14.0	1100	17	CNS016KD	AL106855 Drosophil
C	14	34.8	13.9	967	14	BQ708275	BQ708275 AGENCOURT
C	15	34.8	13.9	1190	14	BQ673485	BQ673485 AGENCOURT
C	16	34.4	13.6	844	17	CNS0052P	AL056652 Drosophil
C	17	34	13.6	767	17	CNS00881	AL051541 Drosophil
C	18	34	13.6	933	12	BF313090	BF313090 601896494
C	19	33.6	13.4	1024	12	BF274287	BF274287 GA_EB002
C	20	33.6	13.4	1138	14	BM915366	BM915366 AGENCOURT
C	21	33.4	13.4	598	10	BS177541	BS177541
C	22	33.4	13.4	604	14	BQ769269	BQ769269 UI-M-F10-
C	23	33.4	13.4	1101	17	CNS017SY	AL108460 Drosophil
C	24	33.4	13.4	1166	14	BM810185	BM810185 AGENCOURT
C	25	33.2	13.3	257	17	TA376C03P	AL497552 T_brucel
C	26	33.2	13.3	524	14	BM762473	BM762473 K-EST0043
C	27	33.2	13.3	539	12	BG324823	BG324823 60243937
C	28	33.2	13.3	564	12	BG481355	BG481355 602528705
C	29	33.2	13.3	589	12	BG501388	BG501388 602547706
C	30	33.2	13.3	712	13	BI890388	BI890388 ZF637-2-C
C	31	33.2	13.3	739	12	BE889325	BE889325 601513213
C	32	33.2	13.3	782	17	CNS01YMS	AL173210 Tetradon
C	33	33.2	13.3	787	12	BF119896	BF119896 601737674
C	34	33.2	13.3	909	14	BQ898865	BQ898865 AGENCOURT
C	35	33.2	13.3	922	12	BG775018	BG775018 602650162
C	36	33.2	13.3	1072	13	BM559867	BM559867 AGENCOURT
C	37	33.2	13.3	1168	13	BM558602	BM558602 AGENCOURT
C	38	33.2	13.3	1350	14	BM912602	BM912602 AGENCOURT
C	39	33.2	13.3	1396	14	BM811251	BM811251 AGENCOURT
C	40	33.2	13.3	1432	14	BM912186	BM912186 AGENCOURT
C	41	33.2	13.3	1432	14	BM912207	BM912207 AGENCOURT
C	42	33	13.2	978	17	AG132188	AG132188 Pan trogl
C	43	32.8	13.1	457	13	BI994587	BI994587 103101360
C	44	32.8	13.1	490	13	BI999790	BI999790 1031078F1
C	45	32.8	13.1	574	12	BE726743	BE726743 894093C09

ALIGNMENTS

RESULT 1
AUI35588
LOCUS AUI35588 PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA
DEFINITION AUI35588 linear EST 02-AUG-2002
sequence.
ACCESSION AUI35588
VERSION AUI35588.1 GI:10996127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 736)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,I., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1332-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; S. & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.


```

FEATURES
  Source
    Location/Qualifiers
      1..736
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="PLACE1002437"
      /clone_lib="PLACE1"
      /issue_type="placenta"
      /note="Vector: pME18SFL3"
BASE COUNT      163 a   199 c   199 g   170 t      5 others
ORIGIN
  Query Match      35.8%; Score 89.4; DB 9; Length 736;
  Best Local Similarity 98.0%; Pred. No. 5.7e-15;
  Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16C GTAATTGGCAGCAGAGTGAGTGGGGCGGGGACCGCGAGCGCGAGCCCTCTCTCTC 219
|||||
DB 1 GTAGTGGCAGCAGAGTGAGTGGGGCGGGGACCGCGAGCGCGAGCCCTCTCTCTC 60
|||||

QY 220 CCGGCTCGCGCAGCGAGCGCGGGCGGGGAGCTC 250
|||||
DB 61 CCGGCTCGCGCAGCGAGCGCGGGCGGGGAGCTC 91
|||||

RESULT 2
LOCUS      AL698654      547 bp      mRNA      linear      EST 21-MAR-2002
DEFINITION DK22P686N12109_21 686 (synonym: hlec3) Homo sapiens cDNA clone
ACCESSION   AL698654
VERSION     AL698654.1 GI:19619194
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 547)
AUTHORS     Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
TITLE       EST (Duesterhoeft, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Duesterhoeft A
MIPS       Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No si sequence available.
            This clone (DKFZp686N12109) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
  source
    Location/Qualifiers
      1..547
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="DKFZp686N12109"
      /clone_lib="586 (synonym: hlec3)"
      /issue_type="human skeletal muscle"
      /dev_stage="adult"
      /lab_host="DH10B"
      /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"
BASE COUNT      112 a   154 c   157 g   124 t
ORIGIN
  Query Match      35.6%; Score 89; DB 9; Length 547;
  Best Local Similarity 100.0%; Pred. No. 6.2e-15;
  Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AATTGGCAGCAGAGTGAGTGGGGCGGGGACCGCGAGCGCGAGCCCTCTCTCTC 221
|||||
DB 1 AATTGGCAGCAGAGTGAGTGGGGCGGGGACCGCGAGCGCGAGCCCTCTCTCTC 60
|||||

RESULT 4
LOCUS      Z44377      292 bp      mRNA      linear      EST 14-NOV-1994
DEFINITION HSC123081 normalized infant brain cDNA Homo sapiens cDNA clone
ACCESSION   Z44377
VERSION     Z44377.1 GI:573506
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 292)

```

AUTHORS Auffray, C., Betar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, P., Miccilli, H., Mariage-Samson, R., Plecu, G., Pouliot, Y., Sebastiani-Kabatchis, C. and Tessier, A.

C-ITILE IMAGE: molecular integration of the analysis of the human genome and its expression

JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

MEDLINE 95277534

COMMENT Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, B960 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160776598
Email: genexpress@genethon.fr

Genexpress_library_id: C; Genexpress_sequence_id: ylc-12b08
Seq primer: (-21)M13_universal.

FEATURES
Location/Qualifiers
1..292
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-12b08"
/clone_lib="normalized infant brain: cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex:Female; dev_stage=3 months old; Isolatemuscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dI) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B.Souares, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 96 g 56 t 3 others

ORIGIN

Query Match 30.0%; Score 75; DB 14; Length 292;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 GTGAGTGGGGCGGACCGGAGCGGAGCGGAGCCCTCTCTCCGGGCTCGGCGAGGG 235
Db 1 GTGAGTGGGGCGGACCGGAGCGGAGCGGAGCCCTCTCTCCGGGCTCGGCGAGGG 60

QY 236 CAGGCGGGGAGCTC 250
Db 61 CAGGCGGGGAGCTC 75

RESULT 5
BC034824 1004 bp mRNA linear HTC 26-JUL-2002

LOCUS BC034824

DEFINITION Homo sapiens, clone IMAGE:4749735, mRNA.

ACCESSION BC034824

VERSION BC034824.1 GI:21961568

KEYWORDS HTC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1004)
Strausberg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.

REMARK

COMMENT

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarstine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 52 Row: 3 Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5915657
This clone has the following problem: incomplete processing.

FEATURES
Location/Qualifiers

1..1004
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749735"
/tissue_type="Skin, squamous cell carcinoma"
/clone_lib="NCI-CCAP_Skn4"
/lab_host="DH-OB"
/note="Vector: pCMV-SPORT6.ccdB"

BASE COUNT 306 a 234 c 230 g 234 t

ORIGIN

Query Match 30.0%; Score 75; DB 11; Length 1004;
Best Local Similarity 100.0%; Pred. No. 6.7e-11;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 GTGAGTGGGGCGGACCGGAGCGGAGCGGAGCCCTCTCTCCGGGCTCGGCGAGGG 235
Db 1 GTGAGTGGGGCGGACCGGAGCGGAGCGGAGCCCTCTCTCCGGGCTCGGCGAGGG 50

QY 236 CAGGCGGGGAGCTC 250

Db 61 CAGGCGGGGAGCTC 75

RESULT 6

BG678861

LOCUS BG678861

DEFINITION 502624760F1 NCI-CCAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5', mRNA sequence.

ACCESSION BG678861

VERSION BG678861.1 GI:13910258

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 998)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10603 row: g column: 16
High quality sequence stop: 860.

FEATURES
Location/Qualifiers

1..998
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/clone="WAGE:4749735"
/clone_lib="NCL_GAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DHI0B (T1 phage-resistant)"
/notes="Organ: skin; Vector: pcwv-sports6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Cligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCL_GAP Library."
BASE COUNT      285 a   233 c   244 g   236 t
ORIGIN
Query Match      24.1%; Score 60.2; DB 12; Length 998;
Best Local Similarity 89.0%; Pred. No. 1.1e-06;
Matches 65; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 178 GAGTGGGCGGACCGGACGAGCGGACCGACCGCTTCTCCCGGGCTGGCGGAGCA 237
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1  GTAGTGGGCGGACCGGACGAGCGGACCGACCGCTTCTCCCGGGCTGGCGGAGCA 60

QY 238 GGGCGGAGCTC 250
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGGCGGAGCTC 73

RESULT 7
LOCUS      CNS006N3/3      912 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14H04 of RpCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL065775      GI:4944655
VERSION     AL065775.1
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
ORGANISM    Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 912)
AUTHORS     Smith, J., Bruley, C.K., Paton, I.R., Dunn, I., Jones, C.F., Windsor, D.,
            Morris, D.R., Law, A.S., Masabanda, J., Sazanov, A., Waddington, D.,
            Fries, R. and Burt, D.W.
            Differences in gene density on chicken macrochromosomes and
            microchromosomes
            Anim. Genet. 31 (2), 96-103 (2000)
MEDLINE     20244064
PUBMED      10782207
REFERENCE   2 (bases 1 to 1863)
AUTHORS     Smith, J., Bruley, C.K., Paton, I.R., Law, A.S., Masabanda, J.,
            Waddington, D., Fries, R. and Burt, D.W.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammos in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RpCI-98 and was constructed by partial,
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
            p1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source            1..912
                        /organism="Drosophila melanogaster"
                        /db_xref="taxon:7227"
                        /clone="BACR14H04"
                        /clone_lib="RpCI-98"
                        /note="end : TET3"
BASE COUNT      178 a   240 c   82 g   146 t   266 others
ORIGIN
Query Match      15.8%; Score 39.6; DB 17; Length 912;
Best Local Similarity 26.7%; Pred. No. 0.74;
Matches 52; Conservative 67; Mismatches 76; Indels 0; Gaps 0;

```

```

QY 50 GAGCGCGGAACGGGCGGAGGAGGAGGACACAGGCTTTTGACCGATAGTAACCTCTG 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 694 GGGGVSQSVVSGGGSSVVVRGVRGGGSGGVSQVGNVGMGGGVGMGNFSAHV 635

QY 110 CGCTGGTGCAGCGGAATCTATAAAGGAAGTAATGTCGGGCAAAACCCCTAATTCGA 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 634 VVGVGVMGMRGVRGVMGMRRRVRGVMGSSMCCSCMCGSGVGRGGSGVSGSC 575

QY 170 CGAGAGTGTAGTGGGCGGACCGGACGAGCGGACCGCTTCTCCCGGGCTCGC 223
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 574 GMSAGSVGMSVGVGAGASRAAGMGCGCGCGCMCCCGCCSCCGCGSGRGA 515

QY 230 GCAGGCGAGGCGGG 244
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 GAVASGCASGGGSR 500

RESULT 8
LOCUS      GGA200110      1863 bp      DNA      linear      GSS 08-JUL-2000
DEFINITION Gallus gallus anonymous sequence from Cosmid mapping to a
            microchromosome (Cosmid 20 - Contig 12), genomic survey sequence.
ACCESSION  AJ231825      GI:3451643
VERSION     AJ231825.1
KEYWORDS    GSS: genome survey sequence.
SOURCE      chicken.
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 1863)
AUTHORS     Smith, J., Bruley, C.K., Paton, I.R., Dunn, I., Jones, C.F., Windsor, D.,
            Morris, D.R., Law, A.S., Masabanda, J., Sazanov, A., Waddington, D.,
            Fries, R. and Burt, D.W.
            Differences in gene density on chicken macrochromosomes and
            microchromosomes
            Anim. Genet. 31 (2), 96-103 (2000)
MEDLINE     20244064
PUBMED      10782207
REFERENCE   2 (bases 1 to 1863)
AUTHORS     Smith, J., Bruley, C.K., Paton, I.R., Law, A.S., Masabanda, J.,
            Waddington, D., Fries, R. and Burt, D.W.
            Direct Submission
            Submitted (12-AUG-1998) Division of Molecular Biology, Roslin
            Institute, Roslin, Midlothian EH25 9PS, Scotland, UK
FEATURES             Location/Qualifiers
     source            1..1863
                        /organism="Gallus gallus"
                        /db_xref="taxon:9031"
                        /chromosome="microchromosome"
BASE COUNT      408 a   392 c   580 g   482 t
ORIGIN
Query Match      15.1%; Score 37.8; DB 17; Length 1863;
Best Local Similarity 54.7%; Pred. No. 2.8;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 17 TGTCTTCTGCTGACTGACTACATACATACAGAGGCGGGAACCGGCGGGGAGAGG 76
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 TGATTACAGAGTGTACTACTACATAGCAGCAGCCACTATTAGCTTTGGCGTAGC 298

QY 77 GAGACACAGGCTTTGACCGATAGTAACCTTCGCGTCGTCGACCGGATCTATAAAG 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 GAAAGAGAAATCTGCAGCTCAITCCAGCCACATCCACCCCTGAGGTGAATATAACAG 358

QY 137 GAACAGTCCCGCAA 153
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 GAATGATCTAGGTAA 375

RESULT 9
LOCUS      CNS016BR      1201 bp      DNA      linear      GSS 26-JUL-1999

```

DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN15K14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL106545
KEYWORDS GSS.
SOURCE AL106545.1 GI:5622456
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Direct Submission
TITLE Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES Location/Qualifiers
source l..1201
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN15K14"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"
BASE COUNT 386 a 182 c 234 g 101 t 298 others
ORIGIN
Query Match 14.9%; Score 37.2; DB 17; Length 1201;
Best Local Similarity 16.0%; Pred. No. 3.7;
Matches 39; Conservative 107; Mismatches 97; Indels 0; Gaps 0;
QY 3 GCCCGGCTCCACGTCCTTTCGTGCTGACTGACTGACTACATAACAGAGCGCGGGAAGC 62
Db 1201 MSASCSVCGGGMGGVGGGCGCGCGGCCMCACAMGCGSAGACMSRSASGMSSS 1142
QY 63 GCGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
Db 1141 CSVSASASSSSVVAARASSANSVSVSVSVSVSVSVSVSVSVSVSVSVSVSVSV 1082
QY 123 CGAATCTATAAAGGAAGTATGTCGCGCAAAACCCGCTAATTCGAGGAGAGTGAAGT 182
Db 1081 MSASCSVCGGGMGGVGGGCGCGGCCMCACAMGCGSAGACMSRSASGMSSS 1022
QY 183 GCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 242
Db 1021 SSSSSRAGASACSVGSSSSSSSSCGAGVGGGCGGSSSSGSCBAMCGGSSSCSS 962
QY 243 GGG 245
Db 961 GGS 959
RESULT 10
CNS012JV
LOCUS 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN08M13 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL101653
KEYWORDS GSS.
SOURCE AL101653.1 GI:5613264
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES Location/Qualifiers
source l..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN08M13"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"
BASE COUNT 293 a 258 c 198 g 32 t 320 others
ORIGIN
Query Match 14.7%; Score 36.8; DB 17; Length 1101;
Best Local Similarity 34.0%; Pred. No. 4.8;
Matches 70; Conservative 42; Mismatches 94; Indels 0; Gaps 0;
QY 38 ACTACATAACAGAGCGCGGGAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97
Db 896 ASACARRAARACACAGARASSHCACGCGGGGGGGGGRMAAMRCACACGAGASCA 955
QY 98 TAGTAACCTCTCGCTCGCTCGAGCGCAATCTATTAAGGAAGTGTCCCGGCAAAACC 157
Db 956 CRGSAACACASACMMCAACACACAAASAMGRMCCVASARACACAGAGCAGACAGMG 1015
QY 158 CCGTATTCGAGGAGAGTGTGAGTGGCGCGGACCCGAGGAGCGGAGCGGAGCGGAG 217
Db 1016 VGARSVGGSCGAGCGARAVAGCRGGAGGMMVSCGMCSCGAGCGGCGGCGGCGG 1075
QY 218 TCCCGGGCTCGCGGCGGAGCGGCGG 243
Db 1076 GCSGGCGGAGCGCGASAGCGRSG 1101
RESULT 11
CNS00GP3
LOCUS 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACR33EL6 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL072367
KEYWORDS GSS.
SOURCE AL072367.1 GI:4952248
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila

Qy	28	GATGCTACTGAAC	TACATAACAGAGCCCGGAAACGGGGCGGAGGAGGACACAGG	87
Ds	973	GVGVMKRVMMVGGGMRMGVGGVGGGGMRGMVGGVYMGVGGMGMVGGVGMVGM		914
Qy	88	CTTTGACCGATAGTAACTCTCGCTCGGTGCAGCGCAATCTATAAAAGAACA	TAGTCCC	147
Ds	913	VVGMGMGMGCGKGYVNGGGMGVYMGHRMGVGVMMGRACMGDMRNVYRMHMMVRG		854
Qy	148	GGCAAAAAACCCGTAATTTCGGAGCGAGAGTGAGTGGGGCGGAGCCGCCAGACCCGAGCC		207
Ds	853	MMGAMGSCSGSVGSGCGCGCMRGVGMGRSKAAGRGMGRVGNCCACAGGGMG		794
Qy	208	GACCCTTCTTCCTCGGCTCGCAGAGGCGAGCGCGGGAGCTC		250
Ds	793	CACACACCGCMSMGAISRGSGGAMCAGVGCMSCGCCRRC		751
RESULT 14				
BQ708275/c				
LOCUS				
DEFINITION BQ708275 NIE_MGC_113 Homo sapiens cDNA clone IMAGE:6301601				
5' mRNA sequence.				
ACCESSION BQ708275.1 GI:21847174				
VERSION				
KEYWORDS EST.				
SOURCE human.				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE 1 (bases 1 to 967)				
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.				
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL Unpublished (1999)				
COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-f@mail.nih.gov Tissue Procurement: Dr. Mark Watson cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Cloned by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2517 row: 1 column: 18 High quality sequence stop: 539.				
FEATURES				
source				
1..967				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:6301601"				
/clone_lib="NIH-MGC_113"				
/lab_host="DH103 (phage-resistant)"				
/note="Organ: spleen; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming, directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G); Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."				
BASE COUNT 92 a 456 c 259 g 129 t 31 others				
ORIGIN				
Query Match 13.9%; Score 34.8; DB 14; Length 967;				
Best Local Similarity 48.9%; Pred. No. 17;				
Matches 87; Conservative 0; Mismatches 91; Indels 3; Gaps 0;				
Qy	63	GGGGGGGAGGAGGAGCAGACAGCTTTGACCATAGTAACCTTCGCTCGGTGAGC		122
Ds	952	GGGGGGGATAGGAGGCGCGGGGGGGGCGAGCAGCGCGGGCGCGGACCCG		893
Qy	123	CGATCTTATAAAGGAAGTCTGCCCGCAAACCCCTAATTCGACGAGATGATG		182
Ds	892	CGGGGAGGGGNNBGCGGGCGCGCGGGGGGGCGCGCACGGCGCGCGCGCG		833

Qy	183	GGGCGCGGACCGCGAGACCGAGCCGACCCCTTCCTCCCGGCTGCGCGACGGCAGGG	240
Db	832	GGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG	775
RESULT 15			
BQ673485			
LOCUS			
DEFINITION	B0673485	1190 bp	linear
	5', mRNA sequence.		
ACCESSION	B0673485		
VERSION	B0673485.1	GI:21784319	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1190)		
AUTHORS	N.H.-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-femail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2405 row: a column: 10 High quality sequence stop: 226. Location/Qualifiers 1. 1190 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6254865" /clone._lib="NIH_MGC_102" /tissue._types="epidermoid carcinoma, cell line" /lab_host="DH10B (prage-resistant)" /note="Organ: salivary gland; Vector: pOTB7; Site:1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	30 a	471 c	581 g
ORIGIN		69 t	39 others
	Query Match	13.9%;	Score 34.8; DB 14; Length 1190;
	Best Local Similarity	48.9%;	Pred. No.18;
	Matches 93; Conservative	0; Mismatches	97; Indels 0; Gaps 0;
Qy	54	CGCGGAACGGGCGGGGAGGAGGACACAGGCTTTGACGATGATGAACCTCTGGCT	113
Db	446	CGGCCCCCGGGGGGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG	505
Qy	114	CGGTGCACGCCGAATCTATAAAGGAACCTAGTCCCGGCAAAACCCGTAATTGGAGCGA	173
Db	506	GGGGCGCGCGCCCCCCCCCGCGGGGCGCGGGGCGCGGGGCGCGCGCGCGCGCGCG	565
Qy	174	GAGTGAGTGGGGCGCGGACCGCGAGCGCGAGCGACCTTCTTCGCGGGCTGCGGCGAG	233
Db	566	CGCGGCGCGGGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG	625
Qy	234	GGCAGGCGCG	243
Db	626	GGCGGGGCGG	635

Tue Feb 4 09:39:53 2003

us-09-596-141c-3_copy_1394_1643.rst

Page 8

Search completed: February 4, 2003, 03:27:02
Job time : 487.757 secs

GenCore version 5.1.3
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: February 3, 2003, 12:56:40 : Search time 11:7338 Seconds
 (without alignments)
 6534.066 Million cell updates/sec
 Title: US-09-596-141C-3_COPY_1394_1643
 Perfect score: 250
 Sequence: 1 gggcccggtccacgtgct.....caggcgagggcgaggagctc 250
 Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0
 Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Patents_NA:*
 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
 5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*
 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35	14.0	7218	1	US-08-232-463-14
C 2	32.6	13.0	3250	4	US-09-222-126B-14
C 3	31.2	12.5	975	1	US-08-671-525B-9
C 4	31.2	12.5	975	1	US-08-672-109B-9
C 5	31.2	12.5	975	1	US-08-842-045-9
C 6	31.2	12.5	975	2	US-08-842-238-9
C 7	31.2	12.5	975	3	US-08-629-335B-9
C 8	31.2	12.5	978	3	US-08-736-281A-17
C 9	31.2	12.5	978	4	US-09-097-231-17
C 10	30	12.0	966	3	US-08-459-318-68
C 11	30	12.0	966	3	US-08-459-318-70
C 12	30	12.0	966	3	US-08-459-318-71
C 13	30	12.0	966	3	US-08-458-609A-68
C 14	30	12.0	966	3	US-08-458-609A-70
C 15	30	12.0	966	3	US-08-458-609A-71
C 16	30	12.0	966	4	US-08-446-872A-68
C 17	30	12.0	966	4	US-08-446-872A-70
C 18	30	12.0	966	4	US-08-446-872A-71
C 19	30	12.0	966	4	US-08-762-227A-68
C 20	30	12.0	966	4	US-08-762-227A-70
C 21	30	12.0	966	4	US-08-762-227A-71
C 22	30	12.0	966	5	PCT-US95-01185-68
C 23	30	12.0	966	5	PCT-US95-01185-70
C 24	30	12.0	966	5	PCT-US95-01185-71
C 25	30	12.0	1817	2	US-08-870-518-8
C 26	30	12.0	81001	4	US-09-750-580-1
C 27	29.8	11.9	5310	3	US-08-870-126-11

C 28 29.8 11.9 8310 4 US-09-445-247-11
 C 29 29.8 11.9 14985 1 US-08-652-972A-6
 C 30 29.8 11.9 14985 5 PCT-US96-06231A-6
 C 31 29.6 11.8 68750 3 US-09-335-409-1
 C 32 29.6 11.8 68750 4 US-09-568-102-1
 C 33 29.6 11.8 68750 4 US-09-567-969-1
 C 34 29.6 11.8 68750 4 US-09-568-480-1
 C 35 29.6 11.8 68750 4 US-09-568-486-1
 C 36 29.6 11.8 68750 4 US-09-568-472-1
 C 37 29.6 11.8 68750 4 US-09-567-899-1
 C 38 29.6 11.8 71989 4 US-09-443-501A-2
 C 39 29.4 11.8 1960 3 US-09-165-240-4
 C 40 29.4 11.8 1960 4 US-09-568-059-4
 C 41 29.4 11.8 4085 3 US-09-165-240-5
 C 42 29.4 11.8 4085 4 US-09-568-059-5
 C 43 29.4 11.8 11219 1 US-07-642-734C-1
 C 44 29.4 11.8 11219 3 US-08-439-009A-1
 C 45 28.6 11.4 4403765 4 US-09-103-840A-2

ALIGNMENTS

RESULT 1
 US-08-232-463-14/C
 : Sequence 14, Application US/08232463
 : Patent No. 5670367
 : GENERAL INFORMATION:
 : APPLICANT: DORNER, F.
 : APPLICANT: SCHEIFLINGER, F.
 : APPLICANT: FALKNER, F. G.
 : TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 : NUMBER OF SEQUENCES: 52
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: 1800 Diagonal Road, Suite 500
 : CITY: Alexandria
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22313-0299
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : OPERATING SYSTEM: IBM PC compatible
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/232,463
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/935,313
 : FILING DATE:
 : APPLICATION NUMBER: EP 91 114 300.6
 : FILING DATE: 26-AUG-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BENT, Stephen A.
 : REGISTRATION NUMBER: 29,758
 : REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703)836-9300
 : TELEFAX: (703)683-4109
 : TELEX: 899149
 : INFORMATION FOR SEQ ID NO: 14:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 7218 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : CLONE: pZcept-F1s
 : US-08-232-463-14

Query Match 14.0% Score 35; DB 1; Length 7218;

RESULT 11
US-08-469-318-70/c
; Sequence 70, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-70

Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 3 GCCCGGCTCCAGCTCTTCTGCTGAGTACGACGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 62
DB 522 GCCCTGGATCTTCTTCTGCTGAGTACGACGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 463
QY 63 GCGCGGGGAGGAGGAGGAGGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 122
DB 462 GCGAGGCGCTAATGGTGTAGCCATGTTGGAGATTATGAGATTCTTTAGACGGAGGAGA 403
QY 123 CGAATCTATAAAGGAAGTACGCTCCCGCAAAACCCCG 160
DB 402 CGGGTTGATAGTAGAGATTGACCGAGCGGTTCCACCG 365

RESULT 12
US-08-469-318-71/c
; Sequence 71, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 196
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,318
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,872
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-318-71
Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 3 GCCCGGCTCCAGCTCTTCTGCTGAGTACGACGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 62
DB 522 GCCCTGGATCTTCTTCTGCTGAGTACGACGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 463
QY 63 GCGCGGGGAGGAGGAGGAGGAGGCTTTGACCGATAGTAACCTCTGCGCTGGTGCAGC 122
DB 462 GCGAGGCGCTAATGGTGTAGCCATGTTGGAGATTATGAGATTCTTTAGACGGAGGAGA 403
QY 123 CGAATCTATAAAGGAAGTACGCTCCCGCAAAACCCCG 160
DB 402 CGGGTTGATAGTAGAGATTGACCGAGCGGTTCCACCG 365

RESULT 13
US-08-468-609A-68/c
; Sequence 68, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairin H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kuman
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (I
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

; MOLECULE TYPE: DNA (genomic)
US-08-468-609A-68

Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78: Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 3 GCCCGGCTCCACGCTGCTTTCGTGAGTCTGACTGACTACATAAACAGAGCGCGGAACG 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 522 GCCCTGGAATCTTCCTCAGTCTCTAAGCACTTGACGAGGAAGCTCTGGGCGAGGAGCT 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 GGGCGGGAGGAGGAGGAGGACAGACGCTTTGACCGAIACTAACCTCTCGCTCGGTGCACC 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 GGCAGGCCCTAATCGGTGAGCCATGTTGGGAGATTATGAGATCTTTAGAGCGGAGAGA 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 CGAATCTATAAAGAACTAGTCCCGGCAAAAACCCCG 160
    || || || || || || || || || || || || || || || || || || || ||
Db 402 CGGCTTATAGATAGAGATTGGACGAGCGGTCCCG 365
    || || || || || || || || || || || || || || || || || || || ||

RESULT 14
US-08-468-609A-70/c
; Sequence 70, Application US/08468609A
; Patent No. 6030612
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Bratford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairé H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olick, Peter O.
; APPLICANT: Paik, Kumman
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (IL-
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60690
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-609A-70

```

Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 3 GCCCGGCTCCAGTGTCTTCCTGAGTGACACTACATCAACAGAGGCCGGGAACG 62
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 GCCTGGAICTTCICACTGCTCTTAGCNCCTTAGCAGGAAGCTCTGGGCAGGAGCT 463
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 GGCGGAGGAGGAGGAGCACAGCGGTTTGACCGATAGTAACCTTCGGTCCGAGC 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 GCGAGGCGCCATAAGTGCTAGCCCATGTTTGSAGATTATGAGATTCTTTAGACGGAGAGA 403
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 CGAICTATAAAGGAACTAGTCCGGCAAAAACCCCG 160
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 CCGGTGATAGTAGAATTGGACCAGACGCTTCAACCG 365

RESULT 15
US-08-468-609A-71/c
; Sequence 71, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maïre H.
; APPLICANT: Easton, Allan W.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Ollis, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (I
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-609A-71

Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Query Match 12.0%; Score 30; DB 3; Length 966;
Best Local Similarity 49.4%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 80; Indels

```
Qy 3 GCCCGGCTCCACGTGCTTCTCTCAGTGAGCTGACACTACATAACAGAGAGCGCGGAGCG 62
Db 522 GCCCTGGATCTTCCTCACTTGGCTTAAGCACTTGAGCAGGAGGCTTGGGGCAGGAGCT 463
Qy 63 GGGCGGGGAGGAGGAGGAGCAGAGGCTTTGACCGATAGTAACTCTGCGCTGCGTGACG 122
Db 462 GGCAGGGCCCTAATGGGTAGCCCATGTTGGAGATTATGAGATTCTTTAGACGGAGGAGA 403
Qy 123 CGAATCTATAAAGGAACTAGTCCCGGCAAAACCCCG 160
Db 402 CGGGTTGATAGTAGAGATTGGACCCAGACGCTTACCCG 365
```

Search completed: February 3, 2003, 21:02:41
Job time : 25.7338 secs


```

; NAME/KEY: misc_feature
; LOCATION: 10945..12946
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 12947..12958
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 13470..13526
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 13641..13752
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 14271..15968
; OTHER INFORMATION: exon 4
; NAME/KEY: misc_feature
; LOCATION: 15969..17969
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1239
; OTHER INFORMATION: 20-828-311 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12347
; OTHER INFORMATION: 17-42-319 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15241
; OTHER INFORMATION: 17-41-250 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 42218
; OTHER INFORMATION: 20-841-149 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 45442
; OTHER INFORMATION: 20-842-115 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 77058
; OTHER INFORMATION: 20-853-415 : polymorphic base C or T
; NAME/KEY: primer_bind
; LOCATION: 929..949
; OTHER INFORMATION: 20-828.pu
; NAME/KEY: primer_bind
; LOCATION: 1357..1377
; OTHER INFORMATION: 20-828.rp complement
; NAME/KEY: primer_bind
; LOCATION: 12029..12050
; OTHER INFORMATION: 17-42.pu
; NAME/KEY: primer_bind
; LOCATION: 2581..12603
; OTHER INFORMATION: 17-42.rp complement
; NAME/KEY: primer_bind
; LOCATION: 14992..15012
; OTHER INFORMATION: 17-41.pu
; NAME/KEY: primer_bind
; LOCATION: 15460..15482
; OTHER INFORMATION: 17-41.rp complement
; NAME/KEY: primer_bind
; LOCATION: 42070..42090
; OTHER INFORMATION: 20-841.pu
; NAME/KEY: primer_bind
; LOCATION: 42572..42591
; OTHER INFORMATION: 20-841.rp complement
; NAME/KEY: primer_bind
; LOCATION: 45328..45347
; OTHER INFORMATION: 20-842.pu
; NAME/KEY: primer_bind
; LOCATION: 45863..45883
; OTHER INFORMATION: 20-842.rp complement
; NAME/KEY: primer_bind
; LOCATION: 76644..76664
; OTHER INFORMATION: 20-853.pu
; NAME/KEY: primer_bind
; LOCATION: 77166..77185
; OTHER INFORMATION: 20-853.rp complement
; NAME/KEY: primer_bind

```

```

; LOCATION: 1220..1238
; OTHER INFORMATION: 20-828-311.mis
; NAME/KEY: primer_bind
; LOCATION: 1240..1258
; OTHER INFORMATION: 20-828-311.mis complement
; NAME/KEY: primer_bind
; LOCATION: 2328..12346
; OTHER INFORMATION: 17-42-319.mis
; NAME/KEY: primer_bind
; LOCATION: 12348..12366
; OTHER INFORMATION: 17-42-319.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15222..15240
; OTHER INFORMATION: 17-41-250.mis
; NAME/KEY: primer_bind
; LOCATION: 15242..15260
; OTHER INFORMATION: 17-41-250.mis complement
; NAME/KEY: primer_bind
; LOCATION: 42199..42217
; OTHER INFORMATION: 20-841-149.mis
; NAME/KEY: primer_bind
; LOCATION: 42219..42237
; OTHER INFORMATION: 20-841-149.mis complement
; NAME/KEY: primer_bind
; LOCATION: 45423..45441
; OTHER INFORMATION: 20-842-115.mis
; NAME/KEY: primer_bind
; LOCATION: 45443..45461
; OTHER INFORMATION: 20-842-115.mis complement
; NAME/KEY: primer_bind
; LOCATION: 77039..77057
; OTHER INFORMATION: 20-853-415.mis
; NAME/KEY: primer_bind
; LOCATION: 77059..77077
; OTHER INFORMATION: 20-853-415.mis complement
; NAME/KEY: misc_binding
; LOCATION: 1227..1251
; OTHER INFORMATION: 20-828-311.probe
; NAME/KEY: misc_binding
; LOCATION: 12335..12359
; OTHER INFORMATION: 17-42-319.probe
; NAME/KEY: misc_binding
; LOCATION: 15229..15253
; OTHER INFORMATION: 17-41-250.probe
; NAME/KEY: misc_binding
; LOCATION: 42206..42230
; OTHER INFORMATION: 20-841-149.probe
; NAME/KEY: misc_binding
; LOCATION: 45430..45454
; OTHER INFORMATION: 20-842-115.probe
; NAME/KEY: misc_binding
; LOCATION: 77046..77070
; OTHER INFORMATION: 20-853-415.probe
; US-09-751-877-1

```

```

Query Match      12.0%; Score 30; DB 10; Length 81001;
Best Local Similarity 61.5%; Pred. No. 15;
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

```

QY 1 GGGCCCCGGCTCCAGTGGCTTCTGCTGAGTGAACACATAACAGAGCGCGGAA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 17479 GGGCTGCTGCTCTCTCTCTCTCGCGCTGATGGCCGGAACAGTGATCAGGGCAGCGG 17420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CGGGCGCGGGGAGGAGGA 78
    ||||| ||||| ||||| |||||
DB 17419 CGGGCGCGGGGAGGAGGA 17402
    ||||| ||||| ||||| |||||

```

```

RESULT 9
US-09-867-701-8750/c
; Sequence 8750: Application US/09867701
; Patent No. US2002013237A;
; GENERAL INFORMATION:

```



```

Db 1465 TAGCATGGAGGGGCTGTGCAGCTGAATGCTCATGTCAGCTGGTGGAGTCTGGAAAT 1524
Qy 192 ATGATGGAGCTGGAGGTGGGAAGAGAGTAGGCTTGGGCGAGCTCTCTCARGCAACCTCA 251
Db 1525 ATGATGGAGCTGGAGGTGGGAAGAGAGTAGGCTTGGGCGAGCTCTCTCARGCAACCTCA 1584
Qy 252 TCTCGGCCAAACATCAGTCAAACTGTGAAGAGTCTAAATGTAATCTGCCCTCAAGGT 311
Db 1585 TTCTGGCCAAACATCAGTCAAACTGTGAAGAGTCTAAATGTAATCTGCCCTCAAGGT 1644
Qy 312 GGCTACAAAGATATCTTTGTCAAGGTAGAGACCTTGTGGCTCCACAGTGCACATTCACGG 371
Db 1645 GGCTACAAAGATATCTTTGTCAAGGTAGAGACCTTGTGGCTCCACAGTGCACATTCACGG 1704
Qy 372 GCGTGGCTT-GCGCTCTCTACGGCTGTCTCTGAGTCTCTATGAATCTCCCTTCAGGGC 430
Db 1705 GCGTGGCTTGGGCTCTCTACGGCTGTCTCTGAGTCTCTATGAAT--CCCTTCAGGGC 1761
Qy 431 AGATTCAATTTAGACTCTTCACAGTTTGACTGAGTTTGGCCAGATAAGGTGACATT 490
Db 1762 AGATTCAATTTAGACTCTTCACAGTTTGACTGAGTTTGGCCAGATAAGGTGACATT 1821
Qy 491 TAGTTTGTGGCTTCATGAATAGCTTAATAATTTAGACATATGCTGTAGGCTTGCATT 550
Db 1822 TAGTTTGTGGCTTCATGAATAGCTTAATAATTTAGAC--ATGGTGTGTAGGCTTGCATT 1879
Qy 551 CCTACTCTGGCTTTTCTTGGCCCTCCAGTGTGTGGGTAGTCTTCTGCCCTCAG 610
Db 1880 CCTACTCTGGCTTTTCTTGGCCCTCCAGTGTGTGGGTAGTCTTCTGCCCTCAG 1938
Qy 611 CCAAGGCAACACAGATAAGTTGGAGGTGTGGAGTGGCTTACATAATTTACACGACTGAA 670
Db 1939 CCAAGGCAACACAGATAAGTTGGAGGTGTGGAGTGGCTTACATAATTTACACGACTGAA 1998
Qy 671 TTCTCTGGCTGCATTCACAAATGTATACAACTAAATACAAAGTCCCTGTGTTTATCAC 730
Db 1999 TTCTCTGGCTGCATTCACAAATGTATACAACTAAATACAAAGTCCCTGTGTTTATCAC 2058
Qy 731 AGGAGGCTGATCAATATAATGAATTAAGAGGGGCTGGCCCATATGCTCTGTGTT 790
Db 2059 AGGAGGCTGATCAATATAATGAATTAAGAGGGGCTGGT-CAATATGCTCTGTGTT 2117
Qy 791 TTG-----TTTGTGTTCTTCTTTTGTGTTTGTGCTCCCTCCCTCAATTA 841
Db 2118 TTGTTTGTGTTTGTGTTCTTCTTTTGTGTTTGTGCTCCCTCCCTCAATTA 2177
Qy 842 TGAAGAGAGCAGTAAGATGTTCTCTGGGCTCTCTGAGGAGCTGGGAGCTCAGGCT 901
Db 2178 TGAAGAGAGCAGTAAGATGTTCTCTGGGCTCTCTGAGGAGCTGGGAGCTCAGGCT 2237
Qy 902 GGAATCTCCAGGAGTAGTCCGCTATCAAAATCAAAATCCAGGTTTGTGGGGGAA 961
Db 2238 GGAATCTCCAGGAGTAGTCCGCTATCAAAATCAAAATCCAGGTTTGTGGGGGAA 2297
Qy 962 AACAAAGCAGCCATACCCAGAGGACTGTCCGCTTCCCTCACCCAGCTTAGGCT 1021
Db 2298 AACAAAGCAGCCATACCCAGAGGACTGTCCGCTTCCCTCACCCAGCTTAGGCT 2357
Qy 1022 TTGAAGGAAACAAAGAGCAAGAAATGATGCGGCTCTGAGGAGATTCAGCTAGA 1081
Db 2358 TTGAAGGAAACAAAGAGCAAGAAATGATGCGGCTCTGAGGAGATTCAGCTAGA 2417
Qy 1082 GCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAAACTAACAAAGGAAAGAAAAATTTG 1141
Db 2418 GCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAAACTAACAAAGGAAAGAAAAATTTG 2477
Qy 1142 CGGAAGCAGGATTTAGAGAGAGCAAAATTCACACTGCTGCCCTTGGCTCCCGGAGCTGG 1201
Db 2478 CGGAAGCAGGATTTAGAGAGAGCAAAATTCACACTGCTGCCCTTGGCTCCCGGAGCTGG 2537
Qy 1202 ACTAGAGAGTCTGGGGCAGCCCGGAGCCGAGCTTCCCGGCTCTTAGCGGGGG 1261

```

```

Db 2538 ACTAGAGAGTCTCGCGCGCAGCCCGAGCCCGAGCTTCCGCGGCTCTAGCGCGGG 2597
Qy 1262 GCCGGGGGGGGAGGGGACGACAGCCGGGACCCCTAAGACACCTGCTGTACCTCCAC 1321
Db 2598 GCCGGGGGGGGAGGGGACGACAGCCGGGACCCCTAAGACACCTGCTGTACCTCCAC 2657
Qy 1322 ----CCCCACCCACCCACCTCCCTCCCACTCCCTAGATGCTGCTGGGGGCTGAACG 1376
Db 2658 CCCACACCCACCCACCCACCTCCCTCCCACTCCCTAGATGCTGCTGGGGGCTGAACG 2717
Qy 1377 TCGCCCGTTTAAAGGGGGGGCCCGGCTCCACGTCTTCTGCTGAGTGAAGTGAACATA 1436
Db 2718 TCGCCCGTTTAAAGGGGGGGCCCGGCTCCACGTCTTCTGCTGAGTGAAGTGAACATA 2777
Qy 1437 TAAACAGAGCCCGGGAAGCGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1496
Db 2778 TAAACAGAGCCCGGGAAGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2837
Qy 1497 CCTCTGCGCTCGGTCCAGCCGAATCTATAAAGGAA 1532
Db 2838 CCTCTGCGCTCGGTCCAGCCGAATCTATAAAGGAA 2873

RESULT 2
US-09-846-456-1
: Sequence 1, Application US/09846456
: Patent No. US20020146792A1
: GENERAL INFORMATION:
: APPLICANT: Rosier, Marie
: APPLICANT: Prades, Catherine
: APPLICANT: Lemoine, Cendrite
: APPLICANT: Naudin, Laurent
: APPLICANT: Denefle, Patrice
: APPLICANT: Duverger, Nicolas
: APPLICANT: Brewer, Bryan
: APPLICANT: Remaley, Alan
: APPLICANT: Fojo, Silvia
: TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
: FILE REFERENCE: 3806.0505
: CURRENT APPLICATION NUMBER: US/09/846.456
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 60/201,280
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 3231
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-846-456-1

Query Match      89.7%;   Score 1374.4;   DB 10;   Length 3231;
Best Local Similarity 97.9%;   Pred. No. 0;
Matches 1504;   Conservative 0;   Mismatches 6;   Indels 26;   Gaps 10;

Qy 16 GGCCTCCACATGCATCTCCAGGCGCTGTTGG--CTCTTCTATGGGCTCTGCTCAGTGT 73
Db 1345 GCGTCCACATGCATCTCCAGGCGCTGCTTGGCGCTCTCTATGCTCTGCTCCTGAGTGT 1404
Qy 74 GATGAGACCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 132
Db 1405 GATGAGACCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1464
Qy 133 TAGCATGGAGGGGCTGT-CAGCTGATGCTGTATGTCAGGTGCTGGGAGTCTCGGAAT 191
Db 1465 TAGCATGGAGGGGCTGTGTCAGCTGAAATGTCATGATGAGGTGCTGGAGTCTTGAAT 1524
Qy 192 ATGATGAGAGCTGGAGGTGGGAAGAGAGTAGGCTTGGGGAGCTCTCTCATGCCACCTCA 251
Db 1525 ATGATGAGAGCTGGAGGTGGGAAGAGAGTAGGCTTGGGGAGCTCTCTCATGCCACCTCA 1584
Qy 252 TTCTGGCCAAACATCAGTCAAACTGTGAAGAGTCTAAATGTAATCTGCCCTCAAGGT 311

```

```
Db 1585 TTTGGCCAAACCTCAGTCAAACTGTGAGAGTCTAAATGTGAATGTGCCCTTCAAGGT 1644
QY 312 GGCTACAAAGATATCTTGTCAAGGTAGGAGACTTGTGGCTCCAGGTGACATTCACAG 371
Db 1645 GGCTACAAAGATATCTTGTCAAGGTAGGAGACTTGTGGCTCCAGGTGACATTCACAG 1704
QY 372 GCGTGCCTTGGCCCTCTAGGGTCTGTCTGAGTCTTCTATGAATCTCCCTTCAGGGC 430
Db 1705 GCGTGCCTTGGCCCTCTAGGGTCTGTCTGAGTCTTCTATGAATCTCCCTTCAGGGC 1761
QY 431 AGATTCATATTAGACTCTTACAGTGTGACCTGAGTGTGGCCAGATAAGGTGACATT 490
Db 1762 AGATTCATATTAGACTCTTACAGTGTGACCTGAGTGTGGCCAGATAAGGTGACATT 1821
QY 490 TAGTTGTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550
Db 1822 TAGTTGTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1879
QY 551 CCTACTCTTGCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 610
Db 1880 CCTACTCTTGCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1938
QY 611 CCAAAGCAACAGATAAGTTGGAGTCTGAGAGGTGATACATAATTTACAGACTGCAG 670
Db 1938 CCAAAGCAACAGATAAGTTGGAGTCTGAGAGGTGATACATAATTTACAGACTGCAG 1998
QY 671 TTCTCTGCTGCTTCTACAAATGATACAAATGATACAAATGATACAAATGATACAA 730
Db 1999 TTCTCTGCTGCTTCTACAAATGATACAAATGATACAAATGATACAAATGATACAA 2058
QY 731 AGGAGGCTGATCAATATATGAAATTTAAAGGGGTGCTGCTGCTGCTGCTGCTGCT 790
Db 2059 AGGAGGCTGATCAATATATGAAATTTAAAGGGGTGCTGCTGCTGCTGCTGCTGCT 2117
QY 791 TCG-----TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 841
Db 2118 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2177
QY 842 TGAAGAGAGCAGTAAGATGTTCTCTCGGTCTCTGAGGAGCTGCGGAGCTCAGGCT 901
Db 2178 TGAAGAGAGCAGTAAGATGTTCTCTCGGTCTCTGAGGAGCTGCGGAGCTCAGGCT 2237
QY 902 GGGAACTCCAAAGCAGTAGTCCCTATCAAAATCAAACTCCAGGTCTTGTGGGGGAA 961
Db 2238 GGGAACTCCAAAGCAGTAGTCCCTATCAAAATCAAACTCCAGGTCTTGTGGGGGAA 2297
QY 962 AACAAAGCAGCCCATTTACCCAGAGGAGTGTCCGCTTCCCTTCCCTTCCCTTCCCT 1021
Db 2298 AACAAAGCAGCCCATTTACCCAGAGGAGTGTCCGCTTCCCTTCCCTTCCCTTCCCT 2357
QY 1022 TTGAAGGAAACAAAGACAGACAAATGATTTGCGCTGCTGAGGAGATTCAGCCTAGA 1081
Db 2358 TTGAAGGAAACAAAGACAGACAAATGATTTGCGCTGCTGAGGAGATTCAGCCTAGA 2417
QY 1082 GCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACTAACAAAGGAAAGGAAAGGAA 1141
Db 2418 GCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACTAACAAAGGAAAGGAAAGGAA 2477
QY 1142 CGGAAGCAGGATTTAGAGGAGCAATTTCCACTGGTGGCTTGGCTGCGGGAGCACTGG 1201
Db 2478 CGGAAGCAGGATTTAGAGGAGCAATTTCCACTGGTGGCTTGGCTGCGGGAGCACTGG 2537
QY 1202 ACTAGAGAGTCTGGGGCAGCCCGAGCCAGCCGCTTCCCGCGGCTTCTAGCGCGGG 1261
Db 2538 ACTAGAGAGTCTGGGGCAGCCCGAGCCAGCCGCTTCCCGCGGCTTCTAGCGCGGG 2597
QY 1262 GCGCGGGGGAGGAGGAGCAGACAGCCGAGCCCTAAGACACCTGCTGTACCCCTCCAC 1321
Db 2598 GCGCGGGGGAGGAGGAGCAGACAGCCGAGCCCTAAGACACCTGCTGTACCCCTCCAC 2657
QY 1322 -----CCCCACCCACCCACTCCCCCAACTCCCTAGATGTGTGTTGGGGGCGGTGAACG 1376
|||||
```

```
Db 2658 CCCACACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 2717
QY 1377 TCGCCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1436
Db 2718 TCGCCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2777
QY 1437 TAAACAGAGGCGCGGGAAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1496
Db 2778 TAAACAGAGGCGCGGGAAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2837
QY 1497 CCTCTGCGCTCGGTGTCAGCCCAATCTATAAAGGAA 1532
Db 2838 CCTCTGCGCTCGGTGTCAGCCCAATCTATAAAGGAA 2873
```

RESULT 3

```
US-09-846-456-3/c
; Sequence 3, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrine
; APPLICANT: Naudin, Laurent
; APPLICANT: Deneffe, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
; TITLE OF INVENTION: Activity and Therapeutic Uses
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846.456
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-3
```

```
Query Match 6.7%; Score 102.2; DB 10; Length 2893;
Best Local Similarity 65.9%; Pred. No. 2e-17;
Matches 164; Conservative 0; Mismatches 83; Indels 2; Gaps 1;
```

```
QY 241 ATGCCACCTCATTTCTGCCAAACTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCIG 300
Db 1820 ATGTACCTTATTTCTGCCAAACTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCIG 1761
QY 301 CCTTCAAGGTGGGTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTTGGCCCTCCACGT 360
Db 1760 CCTTGAAGATTCATAGAAGACTCAGGACACACCCGTAGAAGAGCCCAACGCGCCCTG 1701
QY 361 GCATCTCCAGGCGCTGCTTGGCCCTTCTACGGGTCTGTCTGTAGTCTTCTATGAATCIG 420
Db 1700 GAATGTCACGTGGAGGC--CACAAAGTCTCTACTCTGACAAAGATACCTTTTGGCCAGATG 1643
QY 421 CCTTCAGGCGAGATTCATATTTAGACTCTTACAGTTTACCTGAGTTTGGCCAGATA 480
Db 1642 CTGAGGCGAGATTCATATTTAGACTCTTACAGTTTACCTGAGTTTGGCCAGATG 1583
QY 481 AGGTGACAT 489
Db 1582 AGGTGGCAT 1574
```

RESULT 4

```
US-09-846-456-1/c
; Sequence 1, Application US/09846456
; Patent No. US20020146792A1
```



```
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Catherine
; APPLICANT: Naudin, Laurent
; APPLICANT: Deneffe, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying It
; FILE OF INVENTION: Activity and Therapeutic Uses
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-1

Query Match          6.7%; Score :02.2; DB 10; Length 3231;
Best Local Similarity 65.9%; Pred. No. 2.2e-17;
Matches 164; Conservative 0; Mismatches 89; Indels 2; Gaps 1;

QY 241 ATGCCACTCAITCTGGCCAAACTCAGTCAAACTGTGAAGAGCTCAAAAGTGAATCTG 300
DB 1820 ATGTCACTTATCTGGCCAAACTCAGTCAAACTGTGAAGAGCTCAAAATATGAATCTG 1761
QY 301 CCTTCAGGTGGCTACAAAGTATCTTTGTCAGGTAGGAGACCTGTGGCTCCAGCT 360
DB 1760 CCTCAAGATTCATAGAGACTCAGGACGCCCTGACAGAGGCCCAAGAGGCCCTG 1701
QY 361 GCACATCCAGGCCCTGTGGCTCTCTACGGTCTGTCTGATCTCTATGAATCTC 420
DB 1700 GAAGTGCAGGTGGAGGC--CACAAGGTCTCTACCTGCACAAAGATACCTTTGTAGCCAC 1643
QY 421 CTTTCAGGCGCAATCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATA 480
DB 1642 CTTGAAGGCGAGATTCACATTTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATG 1583
QY 481 AGGTGACAT 489
DB 1582 AGGTGCAAT 1574

RESULT 5
US-09-860-670-255
; Sequence 255, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 32190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-255

Query Match          2.8%; Score 43.4; DB 9; Length 32190;
Best Local Similarity 66.7%; Pred. No. 0.51;
Matches 62; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
```

```
QY 742 TCAATATAAIGAATTAAGAGGGCGCTGCCCATATTGTTCTGTGTTTGTGTTG 801
DB 25616 TAAGAAAAACAACAGAGAGAGCTGATCCCAAGCTACAGGGTTTTCGTTGTTG 25675
QY 802 TTTCTTTTGTGTTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 834
DB 25676 TTTGTTTGTGTTTGTGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25708

RESULT 6
US-09-860-670-260
; Sequence 260, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
; LENGTH: 32249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-260

Query Match          2.8%; Score 43.4; DB 9; Length 32249;
Best Local Similarity 66.7%; Pred. No. 0.51;
Matches 62; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 742 TCAATATAAIGAATTAAGAGGGCGCTGCCCATATTGTTCTGTGTTTGTGTTG 801
DB 18934 TAAGAAAAACAACAGAGAGAGCTGATCCCAAGCTACAGGGTTTTCGTTGTTG 18993
QY 802 TTTCTTTTGTGTTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 834
DB 18994 TTTGTTTGTGTTTGTGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 18926

RESULT 7
US-09-796-692-8332/c
; Sequence 8332, Application US/C9796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,064
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
```

```

; APPLICANT: Van Ruffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 994
; LENGTH: 1974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-834-975-994

Query Match 2.7%; Score 41.6; DB 10; Length 1974;
Best Local Similarity 62.5%; Pred. No. 0.32;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 717 TGTGTTTTATCACAGGAGCTGATCATATATGAATTAAGGGGCTGTCCTCAT 776
Db 221 TTTTITTTGGGGGGGGGGGGTCTTTTATTAAGGGGCCCCCCCCCT 280

Qy 777 ATGTTCTGCTGTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTGT 820
Db 281 TTTTITTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 324

RESULT 10
US-09-764-870-14/c
; Sequence 14, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-14

Query Match 2.7%; Score 41.4; DB 10; Length 2099;
Best Local Similarity 68.7%; Pred. No. 0.38;
Matches 57; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 776 TATGTTCTGCTGTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 835
Db 2094 TTTTITTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 835

Qy 836 AATTTATGAAGAGAGAGTAAG 858
Db 2034 CTTTAAATAAACTAGAGATAAG 2012

RESULT 11
US-09-764-877-3972
; Sequence 3972, Application US/09764877
; Patent No. US2002014743A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

```

; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3972
; LENGTH: 12718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3972

Query Match 2.7%; Score 41; DB 10; Length 12718;

Best Local Similarity 60.2%; Pred. No. 1.3;
Matches 68; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 730 CAGGAGGCTGATCAATAAATGAATAAAGGGGCTGGCCCAATATGTCCTGTT 789

DB 7756 CTGGGTGACAGACAGACCCCTGCTCTRAAAGAAATGCAACATATATCTTTT 7815

QY 790 TTGTTGTTGTTGTTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 842

DB 7816 TTTTCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7868

RESULT 12

US-09-878-574-4299

; Sequence 4299, Application US/09878574

; Patent No. US20020110548A1

; GENERAL INFORMATION:

; APPLICANT: Byrium, Joseph R.

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Thompson, Michael D.

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with:

; FILE OF INVENTION: Plants

; FILE REFERENCE: 38-21(15401)B

; CURRENT APPLICATION NUMBER: US/09/878,574

; CURRENT FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: 09/333,535

; PRIOR FILING DATE: 1999-06-14

; NUMBER OF SEQ ID NOS: 15775

; SEQ ID NO 4299

; LENGTH: 545

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURES:

; NAME/KEY: unsure

; LOCATION: (1)..(545)

; OTHER INFORMATION: unsure at all n locations

; OTHER INFORMATION: Clone ID: LIB3028-013-Q1-B1-H1

US-09-878-574-4299

Query Match

Best Local Similarity 62.1%; Score 40.6; DB 10; Length 545;

Matches 64; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 776 TATTGTCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 835

DB 266 TTTGTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 325

QY 836 AATTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878

DB 326 CCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 368

RESULT 13

US-10-008-118A-19/c

; Sequence 19, Application US/10008118A

; Publication No. US20020187539A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Weng, Zude

; TITLE OF INVENTION: Plant MYB-Related Transcription Factors

; FILE REFERENCE: BB1280 USDB

; CURRENT APPLICATION NUMBER: US/10/008,118A

; CURRENT FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/109,294

; PRIOR FILING DATE: 1998-11-20

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 19

; LENGTH: 1372

; TYPE: DNA

; ORGANISM: Zea mays

US-10-008-118A-19

Query Match

Best Local Similarity 60.6%; Score 40.2; DB 9; Length 1372;

Matches 66; Conservative 3; Mismatches 43; Indels 0; Gaps 0;

QY 776 TATTGTCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 835

DB 1341 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1282

QY 836 AATTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884

DB 1281 TTTTCTTCCAGAGACCCAAAAGATTTCGCCCAATGATCATTTGGCGA 1233

RESULT 14

US-09-443-704-19/c

; Sequence 19, Application US/09443704

; Patent No. US20020066120A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Liu, Zhan-Bin

; APPLICANT: Odell, Joan

; APPLICANT: Rafalski, Antoni

; APPLICANT: Shi, June

; APPLICANT: Weng, Zude

; TITLE OF INVENTION: Plant MYB-Related Transcription Factors

; FILE REFERENCE: BB1280 US NA

; CURRENT APPLICATION NUMBER: US/09/443,704

; CURRENT FILING DATE: 1999-11-19

; EARLIER APPLICATION NUMBER: 60/109,294

; EARLIER FILING DATE: No. US20020066120A1ember 20, 1998

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 19

; LENGTH: 1372

; TYPE: DNA

; ORGANISM: Zea mays

US-09-443-704-19

Query Match

Best Local Similarity 60.6%; Score 40.2; DB 10; Length 1372;

Matches 66; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 776 TATTGTCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 835

DB 1341 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1282

QY 836 AATTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884

DB 1281 TTTTCTTCCAGAGACCCAAAAGATTTCGCCCAATGATCATTTGGCGA 1233

RESULT 15

US-09-764-869-2064/c

; Sequence 2064, Application US/09764869

; Patent No. US20020061521A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007

; CURRENT APPLICATION NUMBER: US/09/764,869

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442

; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 2064
; LENGTH: 16511
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (14362)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-2064

Query Match      2.68; Score 40; DB 10; Length 16511;
Best Local Similarity 51.78; Pred.No. 2.9;
Matches 91; Conservative 0; Mismatches 85; Indels 0; Gaps 0:

QY 580 TGCACCTCCACAAATGCTATACAAACTAATATACAGTCCCTGTGTTTTATATACAGAGGAGGCT 739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1103 TGTACTTTTTCGATAAATTCCTGAGAAATTTTTCCTCAIGTTTATGGCCATAGGAGT 1044
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 740 GATCAATATAATGAATTAAGAGGGGCTGTCCCATATGTCGTGTTTTGTTGTT 799
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1043 TTGTTAATTTTAACTACAAATGTTCAATTTTATTTTTCCTTTTATTTTCA 984
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 800 TGTTCCTTTTGTGTTTGTGGCCICCTTCCTCAATTTATGAAGAGAGAGCACT 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 983 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 928
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: February 3, 2003, 16:28:23
Job time : 365.471 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

R-n on: February 3, 2003, 16:38:05 ; Search time 1614.75 Seconds
(without alignments)
10165.006 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1080_1643

Perfect score: 564

Sequence: 1 gagctctctctccccaatc.....cagggcaggcgaggagctc 564

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rod.*

36: em_hg_sam.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	564	100.0	1167	9	HS252201 Homo sapi
2	564	100.0	1643	6	AX060715 Sequence
3	564	100.0	1643	6	AX060894 Sequence
c	564	100.0	96717	9	AL359182 Human DNA
5	564	100.0	175064	2	AC012230 Homo sapi
6	560.8	99.4	1167	9	AF258623 Homo sapi
7	548	97.2	183999	6	AX092589 Sequence
8	547.4	97.1	3231	6	AX351029 Sequence
9	547.4	97.1	149034	9	AF275948 Homo sapi
10	541	95.9	201144	9	AF287262 Homo sapi
11	519.8	92.2	69570	2	AC021246 Homo sapi
12	456.4	80.9	2893	6	AX351031 Sequence
c	271.6	48.2	69570	2	AC021246 Homo sapi
c	229.8	40.7	90698	2	AC021345 Homo sapi
13	173	30.7	697	9	AF258627 Homo sapi
c	153.4	27.6	186889	2	AL807243 Mus muscu
17	153.2	27.2	278572	10	AF287263 Rattus no
c	153.2	27.2	145833	2	AC125837 Rattus no
19	98	17.4	7260	6	AX253452 Sequence
20	91	16.1	221	6	AX351032 Sequence
21	89.4	15.9	1556	9	AK024328 Homo sapi
22	89	15.8	1750	9	AK022254 Homo sapi
23	75	13.3	9854	6	AX127831 Sequence
24	75	13.3	9854	6	AX139818 Sequence
25	67	11.9	10442	6	AX060713 Sequence
26	67	11.9	10442	6	AX060892 Sequence
27	67	11.9	10442	9	AF285167 Homo sapi
28	60	10.6	10474	6	AX060719 Sequence
29	60	10.6	10474	6	AX060721 Sequence
30	60	10.6	10474	6	AX060898 Sequence
31	60	10.6	10474	6	AX060900 Sequence
32	55.6	9.9	191720	2	AC109982 Rattus no
c	50.4	8.9	125020	9	AF429315 Homo sapi
c	49.6	8.8	96136	2	AC097674 Rattus no
35	49.6	8.8	125020	9	AF429315 Homo sapi
c	49.4	8.8	173127	2	AC131142 Rattus no
c	49	8.7	65780	2	AC111741 Rattus no
c	47.4	8.4	180461	2	AC106563 Rattus no
c	47	8.3	110000	2	AC096324 Rattus no
c	47	8.3	167886	2	AC126316 Rattus no
c	46.6	8.3	53357	2	AC121553 Mus muscu
c	46.4	8.2	184402	2	AC127041 Rattus no
c	46.2	8.2	150305	2	AC114705 Rattus no
c	46.2	8.2	209216	2	AC117126 Rattus no
c	46	8.2	97967	2	AC126984 Rattus no

ALIGNMENTS

RESULT 1
HSA252201
LOCUS
DEFINITION
Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, 5'UTR and promoter region.
ACCESSION
AJ252201
VERSION
ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE
Human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1167)
Porsch-Ozcurumez.W., Langmann.T., Heimerl.S., Borsukova.H., Kaminski.W.E., Drobnik.W., Honer.C., Schumacher.C. and Schmitz.G.
1167 bp DNA linear PRI 10-APR-2001

Pred. No. is the number of results predicted by chance to have a

TITLE The zinc finger protein 202 (2NF202) is a transcriptional repressor of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene expression and a modulator of cellular lipid efflux

J. Biol. Chem. 276 (15), 12427-12433 (2001)

21192304

11279031

2 (bases 1 to 1167)

Porsch-Oezcuernomez,M.K.

Direct Submission

Submitted (05-JAN-2000) Porsch-Oezcuernomez M.K., Institute for

Clinical Chemistry, University of Regensburg,

Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY

Location/Qualifiers

1..1167

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="9q22-31"

/cell_type="leukocyte"

1..1167

/gene="ABC-1"

5..942

/gene="ABC-1"

/function="cholesterol efflux regulatory protein"

896..900

/gene="ABC-1"

1148..1167

/gene="ABC-1"

/function="cholesterol efflux regulatory protein"

BASE COUNT 278 a 315 c 327 g 247 t

ORIGIN

Query Match 100.0%; Score 564; DB 9; Length 1167;

Best Local Similarity 100.0%; Pred. No. 5.8e-125;

Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCTCTCTCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAT 60

|||||

DB 454 GAGCTCTCTCTCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAT 513

|||||

QY 61 TCGGGAAGCAGGATTAGAGGAAGCAAAATCCACTGGTGCCCTTGCGTCCCGGGAACGT 120

|||||

DB 514 TCGGGAAGCAGGATTAGAGGAAGCAAAATCCACTGGTGCCCTTGCGTCCCGGGAACGT 573

|||||

QY 121 GGACTAGAGAGTCTGGCGGCGACGCCGAGCCGAGCGCTCCCGCGGCTTAGCGCGGC 180

|||||

DB 574 GGACTAGAGAGTCTGGCGGCGACGCCGAGCCGAGCGCTCCCGCGGCTTAGCGCGGC 633

|||||

QY 181 GGCGCGGCGGGGGAAGGGGAGCGAGACCGCGGAGACCTAAGACACCTGCTGTACCCCTCC 240

|||||

DB 634 GGCGCGGCGGGGGAAGGGGAGCGAGACCGCGGAGACCTAAGACACCTGCTGTACCCCTCC 693

|||||

QY 241 ACCCGCACCCACCCACCTCCCGCCACCTCCCTAGATGTGTCTGTGGCGGCTGAACGTCG 300

|||||

DB 694 ACCCGCACCCACCCACCTCCCGCCACCTCCCTAGATGTGTCTGTGGCGGCTGAACGTCG 753

|||||

QY 301 CCGGTTTAAGGGCGGCGCGGCTCCACGTGCTTCTGTGTGAGTGAAGTGAAGTGAAGTGA 360

|||||

DB 754 CCGGTTTAAGGGCGGCGCGGCTCCACGTGCTTCTGTGTGAGTGAAGTGAAGTGAAGTGA 813

|||||

QY 361 ACAGAGCCCGGAACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

|||||

DB 814 ACAGAGCCCGGAACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 873

|||||

QY 421 CTGGCTCGGTGAGCGCGAATCTATAAAGAACTAGTCCCGCAAAACCCCGGTATTG 480

|||||

DB 874 CTGGCTCGGTGAGCGCGAATCTATAAAGAACTAGTCCCGCAAAACCCCGGTATTG 933

|||||

QY 481 CGAGCGAGAGTGAAGTGGGGCGGAGCCGAGAGCCGAGCGAGCCCTTCTCTCCCGGCT 540

|||||

DB 934 CGAGCGAGAGTGAAGTGGGGCGGAGCCGAGAGCCGAGCGAGCCCTTCTCTCCCGGCT 993

|||||

QY 541 CGGCGAGGCGAGGCGGGGAGCTC 564

Db 994 GCGGAGGCGAGGCGGGGAGCTC 1017

RESULT 2

AX060715

LOCUS

AX060715

Sequence 3 from Patent WO0078972.

AX060715

AX060715.1

GI:12405104

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Location/Qualifiers

1..1643

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t

ORIGIN

Query Match 100.0%; Score 564; DB 6; Length 1643;

Best Local Similarity 100.0%; Pred. No. 5.8e-125;

Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCTCTCTCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAT 60

|||||

DB 1080 GAGCTCTCTCTCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAT 1139

|||||

QY 61 TCGGGAAGCAGGATTAGAGGAAGCAAAATCCACTGGTGCCCTTGCGTCCCGGGAACGT 120

|||||

DB 1140 TCGGGAAGCAGGATTAGAGGAAGCAAAATCCACTGGTGCCCTTGCGTCCCGGGAACGT 1199

|||||

QY 121 GGACTAGAGAGTCTGGCGGCGACGCCGAGCCGAGCGCTCCCGCGGCTTAGCGCGGC 180

|||||

DB 1200 GGACTAGAGAGTCTGGCGGCGACGCCGAGCCGAGCGCTCCCGCGGCTTAGCGCGGC 1259

|||||

QY 181 GGCGCGGCGGGGGAAGGGGAGCGAGACCGCGGAGACCTAAGACACCTGCTGTACCCCTCC 240

|||||

DB 1260 GGCGCGGCGGGGGAAGGGGAGCGAGACCGCGGAGACCTAAGACACCTGCTGTACCCCTCC 1319

|||||

QY 241 ACCCGCACCCACCCACCTCCCGCCACCTCCCTAGATGTGTCTGTGGCGGCTGAACGTCG 300

|||||

DB 1320 ACCCGCACCCACCCACCTCCCGCCACCTCCCTAGATGTGTCTGTGGCGGCTGAACGTCG 1379

|||||

QY 301 CCGGTTTAAGGGCGGCGCGGCTCCACGTGCTTCTGTGTGAGTGAAGTGAAGTGAAGTGA 360

|||||

DB 1380 CCGGTTTAAGGGCGGCGCGGCTCCACGTGCTTCTGTGTGAGTGAAGTGAAGTGAAGTGA 1439

|||||

QY 361 ACAGAGCCCGGAACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

|||||

DB 1440 ACAGAGCCCGGAACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1499

|||||

QY 421 CTGGCTCGGTGAGCGCGAATCTATAAAGAACTAGTCCCGCAAAACCCCGGTATTG 480

|||||

DB 1500 CTGGCTCGGTGAGCGCGAATCTATAAAGAACTAGTCCCGCAAAACCCCGGTATTG 1559

|||||

QY 481 CGAGCGAGAGTGAAGTGGGGCGGAGCCGAGAGCCGAGCGAGCCCTTCTCTCCCGGCT 540

|||||

DB 1560 CGAGCGAGAGTGAAGTGGGGCGGAGCCGAGAGCCGAGCGAGCCCTTCTCTCCCGGCT 1619

|||||

QY 541 GCGGCGAGGCGAGGCGGGGAGCTC 564

|||||

DB 1620 GCGGCGAGGCGAGGCGGGGAGCTC 1643

|||||

RESULT 3

```

AX060894
LOCUS AX060894 1643 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078971.
ACCESSION AX060894
VERSION AX060894.1 GI:12406271
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE ATP binding cassette transporter protein abcl1 polypeptides
JOURNAL Patent: WO 0078971-A 3 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
Location/Qualifiers
source
1..1643
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 370 a 413 c 457 g 403 t
ORIGIN
Query Match 100.0%; Score 564; DB 6; Length 1643;
Best Local Similarity 100.0%; Pred. No. 5.8e-125;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0:
QY 1 GAGCTCTCTCCCGCAATCCCTCCCTCCGCTGAGGAAGCAAACTACAAAGAAAAAAT 60
Db 1080 GAGCTCTCTCCCGCAATCCCTCCCTCCGCTGAGGAAGCAAACTACAAAGAAAAAAT 1139
QY 61 TCGGAAAGCAGGATTAGAGGAAGCAAACTCCACTGCTGCCCTTGCTCCGCGGAACGT 120
Db 1140 TCGGAAAGCAGGATTAGAGGAAGCAAACTCCACTGCTGCCCTTGCTCCGCGGAACGT 1199
QY 121 GGACTAGAGAGTTCGCGCGCAGCCGCGAGCCAGCCAGCCCTCCGCGCGCTTTAGCCCGGC 180
Db 1200 GGACTAGAGAGTTCGCGCGCAGCCGCGAGCCAGCCAGCCCTCCGCGCGCTTTAGCCCGGC 1259
QY 181 GGCCCGGGCGGGAGGGAGGACGACACCGCGGACCCCTAGACACCTGCTGTACCTCC 24C
Db 1260 GGCCCGGGCGGGAGGGAGGACGACACCGCGGACCCCTAGACACCTGCTGTACCTCC 1319
QY 241 ACCCCACCCACCCACCTCCCGCCCACTCCCTAGATGTGTGCGGGCGGTGAACCTCG 300
Db 1320 ACCCCACCCACCCACCTCCCGCCCACTCCCTAGATGTGTGCGGGCGGTGAACCTCG 1379
QY 301 CCGCTTTAAGGGCGGGCGCGCGCTCCACGTGCTTTCTGTGTAGTGACTGAACATACATAA 360
Db 1380 CCGCTTTAAGGGCGGGCGCGCGCTCCACGTGCTTTCTGTGTAGTGACTGAACATACATAA 1439
QY 361 ACAGAGCGGGAGACGGCGGGGAGGAGGAGGAGGACACAGGCTTGACCGATACCT 420
Db 1440 ACAGAGCGGGAGACGGCGGGGAGGAGGAGGAGGAGGAGGCTTTGACCGATACCT 1499
QY 421 CTGGCTCGGTGACGCGCAATCTATAAGGAAGTACTGCTCCGCGCAAAACCCCGTAATTG 480
Db 1500 CTGGCTCGGTGACGCGCAATCTATAAGGAAGTACTGCTCCGCGCAAAACCCCGTAATTG 1559
QY 481 CGACGAGAGTGTAGTGGGCGGGAGCCCGGACGCGGAGCGGAGCCCTTCTCCCGGGCT 540
Db 1560 CGACGAGAGTGTAGTGGGCGGGAGCCCGGAGCGGAGCGGAGCCCTTCTCCCGGGCT 1619
QY 541 CGCGCAGGCGGCGGGGAGCTC 564
Db 1620 CGCGCAGGCGGCGGGGAGCTC 1643
RESULT 4
LOCUS AL359182 96717 bp DNA linear PRI 11-JAN-2002
DEFINITION Human DNA sequence from clone RP11-217B7 on chromosome 9, complete
sequence.
ACCESSION AL359182

```

```

VERSION AL359182.20 GI:19151453
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 96717)
REFERENCE
Skuce, C.
Direct Submission
Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 15, 2002 this sequence version replaced gi:18121468.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.; EMBL; SW.;
SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C-elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-217B7 is from the library RP11-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.choi.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-217B7. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-217B7 is at 96717 in this
sequence. The true left end of clone RP11-122F10 is at 72980 in
this sequence. The true right end of clone RP11-3LJ20 is at 2000 in
this sequence.
FEATURES
source Location/Qualifiers
1..96717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-217B7"
/clone_lib="XPC1-11.1"
3238..3278
/notes="Sequence from AP275948 sequenced by National
Institutes of Health, National Heart, Lung and Blood
Institute, Bethesda, MD 20892, USA."
84249..84273
/notes="Sequence from overlapping clone RP11-122F10
(AC026643). Assembly confirmed by restriction digest."
92050..92163
/notes="Sequence from reads from a short insert library
derived from a single pUC clone. Restriction digest data
confirm the assembly."
92411..92557
/notes="Sequence from reads from a short insert library
derived from a single pUC clone. Restriction digest data
confirm the assembly."
BASE COUNT 27673 a 21138 c 20380 g 27526 t
ORIGIN
Query Match 100.0%; Score 564; DB 9; Length 96717;
Best Local Similarity 100.0%; Pred. No. 6.2e-125;

```


AF258623S1 1167 bp DNA linear PRI 23-JUN-2000

LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene, promoter and exon 1.

ACCESSION AF258623

VERSION AF258623.2

KEYWORDS GI:8677405

SEGMENT 1 of 4

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1167)

AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Acuizerat,B.E., Fielding,C.J. and Kane,J.P.

TITLE Analysis of hABCA1 gene 5' end: additional peptide sequence, promoter region, and four polymorphisms

JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press

REFERENCE 2 (bases 224 to 1167)

AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Acuizerat,B.E., Fielding,C.J. and Kane,J.P.

TITLE Direct Submission

JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA

REFERENCE 3 (bases 1 to 1167)

AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Acuizerat,B.E., Fielding,C.J. and Kane,J.P.

TITLE Direct Submission

JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA

REMARK Sequence update by submitter

COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.

FEATURES

source 1..1167

 /organism="Homo sapiens"

 /db_xref="taxon:9606"

 /chromosome="9"

 /map="9q31"

 224..844

 /feature="ABCA1"

 845..1147

 /feature="ABCA1"

 /number=1

BASE COUNT 278 a 313 c 328 g 248 t

ORIGIN

Query Match 99.4%; Score 560.8; DB 9; Length 1167;

Best Local Similarity 99.6%; Pred. No. 3.4e-124;

Matches 562; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY 1 GAGCTCTCTCTCCCCCAATCCCTCCCTCCGCTGAGGAACCTAACAAAGGAAAAAAT 60

Db 454 GAGCTCTCTCTCCCCCAATCCCTCCCTCCGCTGAGGAACCTAACAAAGGAAAAAAT 513

QY 61 TCGCGAAGCAGGATTATGAGGAAGCAAAATCCACIGTGCCTTGGCTGCCGGGAACGT 120

Db 514 TCGCGAAGCAGGATTATGAGGAAGCAAAATCCACIGTGCCTTGGCTGCCGGGAACGT 573

QY 121 GGACTAGAGATCTCGCGCGCAGCCCGAGCCCGAGCCGCTTCGCGCGCTCTTAGCCCGGC 180

Db 574 GGACTAGAGATCTCGCGCGCAGCCCGAGCCCGAGCCGCTTCGCGCGCTCTTAGCCCGGC 633

QY 181 GGGCCCGGGCGGGAAGGGAGCCGAGACCCGGGACCTTAGACACCTCTGTACCTCC 240

Db 634 GGGCCCGGGCGGGAAGGGAGCCGAGACCCGGGACCTTAGACACCTCTGTACCTCC 693

QY 241 ACCCCCAACCCCAACCCACTCCCTCCCACTCCCTAGATGTGTGCGGCGGTGAACGTCG 300

Db 694 ACCCCCAACCCCAACCCACTCCCTCCCACTCCCTAGATGTGTGCGGCGGTGAACGTCG 753

QY 301 CCGCTTTAAGGGCGGGCCCGCGCTCCACCTGCTTCTGCTGAGTGAAGTGAACATAC 360

Db	28672	CAATAACAGAGCGCGGAGACGGGGCGGGAGGAGGAGACACAGGCTTTGACCGATAGT	28731
Qy	416	AACCTCTGCGCTCGGTCCAGCCGCAATCTATAAAGGAACCTAGTCCCGGCAAAACCCCGT	475
Db	28732	AACCTCTGCGCTCGGTCCAGCCGCAATCTATAAAGGAACCTAGTCCCGGCAAAACCCCGT	28791
Qy	476	AATGCGAGGCGAGTGCAGTGGGCGGGGACCGCCGACAGCGCGGAGCCGACCTTCCTCC	535
Db	28792	AATGCGAGGCGAGTGCAGTGGGCGGGGACCGCCGACAGCGCGGAGCCGACCTTCCTCC	28851
Qy	536	GGGCTGGCGGAGGSCAGGGGGGGGAGCTC	564
Db	28852	GGGCTGGCGGAGGSCAGGGGGGGGAGCTC	28880
<p>RESULT 8</p> <p>AX351029 3231 bp DNA linear PAT 06-FEB-2002</p> <p>LOCUS AX351029</p> <p>DEFINITION Sequence 1 from Patent WO0183746.</p> <p>ACCESSION AX351029</p> <p>VERSION AX351029.1 GI:18616395</p> <p>KEYWORDS human.</p> <p>SOURCE human.</p> <p>ORGANISM Homo sapiens</p> <p>EUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>REFERENCE 1</p> <p>AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.</p> <p>TITLE Regulatory nucleic acid sequences of the abcl gene</p> <p>JOURNAL Patent: WO 0183746-A 1 08-NOV-2001;</p> <p>AVENTIS PHARMA S.A. (FR)</p> <p>FEATURES</p> <p>Location/Qualifiers</p> <p>1..3231</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p>			
BASE COUNT	809 a	773 c	876 g
ORIGIN	97.1%; Score 547.4; DB 6; Length 3231;		
Query Match	Best Local Similarity 96.9%; Pred. No. 5.7e-121;		
Matches	563; Conservative 0; Mismatches 1; Indels 5; Gaps 1;		
Qy	1	GAGCTCTCTCTCCCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAAT	60
Db	2416	GAGCTCTCTCTCCCCCAATCCCTCCCTCGGCTGAGGAACCTAACAAAGGAAAAAAAT	2475
Qy	61	TGCGGAAGCAGGATTATAGAGAAAGCAAAATCCACTGCGCCCTTGCTCCGCGGAACGT	120
Db	2476	TGCGGAAGCAGGATTATAGAGAAAGCAAAATCCACTGCGCCCTTGCTCCGCGGAACGT	2535
Qy	121	GGACTAGAGATCTGGCGGCGAGCCCCGAGCCAGCCCTTCCCGCGGCTTTAGGCGGCG	180
Db	2536	GGACTAGAGATCTGGCGGCGAGCCCCGAGCCAGCCCTTCCCGCGGCTTTAGGCGGCG	2595
Qy	181	GGGCGCGGGGGGGAAGGGAGCAGACCGGACCCCTTAAGACACCTGCTGTACCCCTCC	240
Db	2596	GGGCGCGGGGGGGAAGGGAGCAGACCCGAGCCCTTAAGACACCTGCTGTACCCCTCC	2655
Qy	241	AC-----CCCCACCCCAACCTCCGCCCAACTCCCTAGATGTGTGTTGGCGGCGCTGAA	295
Db	2656	ACCCCAACCCCAACCCCAACCTCCGCCCAACTCCCTAGATGTGTGTTGGCGGCGCTGAA	2715
Qy	296	CGTGGCCGGTTTAAAGGGGGGGCCCGGCTCCAGTCGCTTCTGCTGAGTGACTGAACTA	355
Db	2716	CGTGGCCGGTTTAAAGGGGGGGCCCGGCTCCAGTCGCTTCTGCTGAGTGACTGAACTA	2775
Qy	356	CATAACAGAGGCGGGAAGCGGGCGGGAGGAGGAGACACAGGCTTTCACCGATAGT	415
Db	2776	CATAACAGAGGCGGGAAGCGGGCGGGAGGAGGAGACACAGGCTTTCACCGATAGT	2835
Qy	416	AACCTCTGCGCTGCGTGCACCCGCAATCTATAAAGGAACCTAGTCCCGGCAAAACCCCGT	475

Db 2836 AACCTCTCGCTGGTGGAGCCGAATCTATAAAGGAAGTACTCCCGGGCAAAACCCCT 2895
|||||
Qy 476 AATTGCGAGCGAGAGTGGTGGCGGGGACCGCGAGAGCCGAGCCGACCGCTTCTCTCC 535
|||||
Db 2896 AATTGCGAGCGAGAGTGGTGGCGGGGACCGCGAGAGCCGAGCCGACCGCTTCTCTCC 2955
|||||
Qy 536 GGGCTGGCGGAGGCGAGGCGGGGAGCIC 564
|||||
Db 2956 GGGCTGGCGGAGGCGAGGCGGGGAGCIC 2984
|||||

RESULT 9
AF275948 149034 bp DNA linear PRI 17-JUL-2000
LOCUS Homo sapiens ABCA1 (ABCA1) gene, complete cds.
DEFINITION
ACCESSION AF275948
VERSION AF275948.1 GI:9247085
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 149034)
Santamarina-Fojo,S., Peterson,K., Knapper,C., Giu,X., Freeman,L.,
Cheng,J.F., Osorio,J., Remaley,A., Yang,X.P., Haudenschild,C.,
Prades,C., Chimini,G., Blackmon,E., Francois,T., Duverger,N.,
Rubin,E.M., Rosier,M., Deneffe,P., Fredrickson,D.S. and Brewer,H.B.
Jr.
Complete genomic sequence of the human ABCA1 gene: analysis of the
human and mouse ATP-binding cassette A promoter
Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
20345099
PUBMED 10884428
2 (bases 1 to 149034)
Santamarina-Fojo,S., Peterson,K.M., Knapper,C.L., Freeman,L.A.,
Remaley,A.I., Yang,X.-P., Haudenschild,C.C., Blackmon,E.E.,
Francois,T.L. and Brewer,H.B. Jr.
Direct Submission
Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA
Location/Qualifiers
1..149034
/organism="Homo sapiens"
/db_xref="taxon:9606"
1454..148034
/gene="ABCA1"
join(1454..1674,25831..25989,40385..40478,45012..45153,
46423..46541,67729..67850,70831..71007,83960..84052,
89010..89250,91962..92101,92433..92549,96758..96955,
97702..97907,98428..98604,100391..100613,102360..102581,
103642..103846,104951..105064,105862..107033,
108023..108154,109460..109602,109806..109943,
110446..110866,112134..112196,113183..113385,
115053..115101,115297..115410,116806..116954,
118604..118728,119964..120062,123094..123283,
124804..124898,126208..126240,127363..127468,
128943..129017,129539..129708,130939..131116,
133114..133229,133341..133485,134526..134649,
135736..135885,136130..136250,137038..137100,
138008..138114,140469..140610,140982..141116,
142060..142163,142646..142738,143397..143640,
144581..148034)
/gene="ABCA1"
/product="ABCA1"
join(1454..1674,25831..25923)
/gene="ABCA1"
2406..2678
/rpt_family="Alu"
repeat_region 4240..4509
/rpt_family="Alu"
repeat_region 4998..5270

```
repeat_region /rpt_family="Alu" 5563. .5839
repeat_region /rpt_family="Alu" 6615. .6877
repeat_region /rpt_family="Alu" 8800. .9046
repeat_region /rpt_family="Alu" 9998. .10279
repeat_region /rpt_family="Alu" 11859. .12133
repeat_region /rpt_family="Alu" 12810. .12902
repeat_region /rpt_family="Alu" 15220. .15399
repeat_region /rpt_family="Alu" 15508. .15787
repeat_region /rpt_family="Alu" 18600. .18860
repeat_region /rpt_family="Alu" 20950. .21206
repeat_region /rpt_family="Alu" 25036. .25314
repeat_region /rpt_family="Alu" 25036. .25314
CDS
join(25924. .25989,40385. .40478,45012. .45153,46423. .46541,
67729. .67850,70831. .71007,83960. .84052,89010. .89250,
91962. .92101,92433. .92549,95758. .96955,97702. .97907,
98428. .98604,100391. .100613,102360. .102581,103642. .103846,
104951. .105064,106862. .107033,108023. .108154,
109460. .109602,109806. .109943,110646. .110866,
112124. .112196,113183. .113385,115053. .115101,
115297. .115410,116806. .116954,118604. .118728,
119964. .120062,123094. .123283,124804. .124898,
126208. .126240,127363. .127458,128943. .129017,
129539. .129708,130939. .131116,133114. .133229,
133341. .133485,134326. .134649,135736. .135865,
136130. .136250,137038. .137100,138008. .138114,
140469. .140610,140982. .141116,142080. .142163,
142646. .142738,143397. .143640,144581. .144721)
/gene="ABCA1"
/codon_start=1
/product="ABCA1"
/protein_id="AAF96276.1"
/db_xref="GI:9247086"
/translation="MACPOLRLLLWKNLTFRRRTQCLLEVAWPLFIFILISVRL
SYPIEHECHFPKAMPESAGTLFWQGIICNANPCFRPYTPGEAPGVYGNFKSIV
ARLPSDARRLLYSOKOTSNKDMKVLRTLQIQIKSSNLKLODPLVDNFTPSGLYH
NLSLPKSTVDKMLRADVLHKVFLQGYQLHTSLCNGSKSEMIQLGDQVSELCGLP
BEKLAARVLRASWDILKPLIRIINSTSPPSKELAEAKYLLHSLGTLAQELFSMR
SWDMRQEVFLINVNSSSTQIQAVSRIVCGPEGGGLKIKSLRWEDNNYKALF
GGNTEDEDAETFDNSTPYCNDLKNLLESSPLSRIIWKALKPLLVGLKIYTPDPAI
QQLDLQWTADQIVAVFLAKHPEDVQSSNGSVYTWREAFNETNOAIRTISRMECVNLN
RQVMAEYKTFQELAEVFDLEGWEEELSPKIWTPEMSQEMDLVRLDSDRNDPFWF
KLEPIATEVWLINKSMELLDERKFWAGIVFTGTPGSELPHVVKYKIRMDIDNVERT
NKIDGYDPCPRADPREDMRYVWGFEAYLODAVVEQALIRVLIGTEKKIGYMQMPY
PCYVDDIPLRVNSSSMPLFMTLAWIYSAVVIKGIIVIEKEARKRETRIMGLDNLNLW
FSWFISSLIPLLVLAGLLVILKUGNLPLPSDPSVEVFLSVFAVTLIQCLFLSTLP
SRNALAACGIIITFYLLPYLVCAWQYVFTLKIFAXLLSPVAFGCGEYFALFE
EQIGVQNDLNFESQSDSDAGLSDHSDIETLIDVSAISNLIRKHSFARLVPEIDICHEL
YFPCYKAWFGESDEKSHPGSNOKRISIEICMEEPHLLKLGVSIONLAKVYRDMKV
AVDGLALNFYEGQITSPILGNGAGKTMTMSLTGLFPPTSGTAVILGKDISENSTR
QNLGVCHOHNVLFMLAVEEHIWYALKGLSEKHVKAEMQEMALDVLPLSPKLSKT
SOLSGMOKUSALAFVGGKSVILDEPTAGVDYFYSRGVIELLKYRQCTIILST
HMDDEAVLGRIAIISHGKCCVGSLLFLKNQGTGYLFLVKRDVSSSSRNS
STVSLYKDESDVSGSDAGLSDHSDIETLIDVSAISNLIRKHSFARLVPEIDICHEL
TYVLPYAAKGAQAVELFPEIDDLSDLGISYVGISETPLEEILKVAEPPSGVDAETS
DGLTPARNRRAFGCKOSCI RPTTEDDAADNDSDIDPESRETLDSGMDKGSYQVK
GWLTKQCFVALLKRGJIAARRSGFFAQIVLPAVFVCIALVFSLLVPPGKYPSE
LQPMWYEQYITFVSNDAPEDTGTLELNLATKDPGFCRMEGNIIPDTPCGAAGEEW
TTAPVPTIMDLFQNGWNTMONSPACQSSDKIKMLKPLVPGAGGLPPLPPQKQNTA
DILQDLGRNLSVLYKTYVQI IAKSLNKNWNEFRYGGFSLGVSNTQALPPSQEVA
DAXKOMKHLKLADESSADRLNSLRMIGLDTENNVKWNKNGHAISSFLNLIN
NALIRANLQKGENSHGCIATFNPLNLTKOOLSEVAXMTTSDVNLVSICVIFAMSIV
PASFVFLVLIQENSKAKHLQFISGVKPVIIWLSNEFVMDCNKYVPATILVITIFCFQQ
```

```
repeat_region /rpt_family="Alu" 29916. .30180
repeat_region /rpt_family="Alu" 31050. .31314
repeat_region /rpt_family="Alu" 31692. .31973
repeat_region /rpt_family="Alu" 36167. .36440
repeat_region /rpt_family="Alu" 37617. .37892
repeat_region /rpt_family="Alu" 37999. .38166
repeat_region /rpt_family="Alu" 38172. .38269
repeat_region /rpt_family="Alu" 38995. .39271
repeat_region /rpt_family="Alu" 41347. .41565
STS
/gene="ABCA1"
/note="A009X28"
41593. .41835
/rpt_family="Alu" 41657. .41679
/gene="ABCA1"
/note="A009X28"
42320. .42584
/rpt_family="Alu" 44349. .44526
/rpt_family="Alu" 48473. .48669
/rpt_family="Alu" 50547. .50614
/rpt_family="Alu" 54285. .54418
/rpt_family="Alu" 54849. .54988
/rpt_family="Alu" 55029. .55226
/rpt_family="Alu" 55297. .55344
/rpt_family="Alu" 58501. .58935
/note="LTR"
/rpt_family="HERV" 51246. .61489
/rpt_family="Alu" 62812. .63077
/rpt_family="Alu" 64082. .64306
/rpt_family="Alu" 67123. .67382
/rpt_family="Alu" 68499. .68731
/rpt_family="Alu" 69481. .69760
/rpt_family="Alu" 71998. .72233
/rpt_family="Alu" 72244. .72501
Query Match 97.1%; Score 547.4; DB 9; Length 149034;
Best Local Similarity 98.9%; Pred. No. 6e-121; Indels 5; Gaps 1;
Matches 563; Conservative 0; Mismatches 1;
```



```

DEFINITION Homo sapiens clone RP11-IN10, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021246
VERSION AC021246.2 GI:9119882
KEYWORDS HPG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 69570)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome, clone RP11-IN10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 69570)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boquelavkiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
Choelel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Deaileiro,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kan,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGuirk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using repeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: L_N10
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 872 971: contig of 871 bp in length
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3745 3845: gap of 100 bp
* 3845 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 868 bp in length
* 7510 7609: gap of 100 bp

```

```

* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length

```

```
* 42925 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48531: contig of 852 bp in length
* 48532 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 851 bp in length
* 55230 55329: gap of 100 bp
* 55330 55197: contig of 868 bp in length
* 55198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp
```

```
Query Match 92.2%; Score 519.8; DB 2; Length 69570;
Best Local Similarity 99.3%; Pred. No. 2.5e-114;
Matches 543; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 20 CCGCTCCCGCTGAGGAACCTAACAA-GGAAAAAAATTCGGAAGCAGGATTTA 78
DB 41109 CCTCCCTCCGCTGAGGAACCTAACAAAGGAAAAAAATTCGGAAGCAGGATTTA 41168

QY 79 GAGGAAGCAATTCCTACTGGTCCCTTGGCTCGCGGGAACGTGACTAGAGATCTCGCG 138
DB 41169 GAGGAAGCAATTCCTACTGGTCCCTTGGCTCGCGGGAACGTGACTAGAGATCTCGCG 41228

QY 139 CGAGCCCGGAGCCCGAGCGCTTCCCGCGGCTTTTAGCCCGGCGGCGCCCGGGGGGAAG 198
DB 41229 CGAGCCCGGAGCCCGAGCGCTTCCCGCGGCTTTTAGCCCGGCGGCGCCCGGGGGGAAG 41288

QY 199 GGGACGACGACCGCGACCCCTAGACACCTGTGTACCTCCACCCACCCACCCACC 258
DB 41289 GGGACGACGACCGCGACCCCTAGACACCTGTGTACCTCCACCCACCCACCCACC 41348
```

```
QY 259 TCCGCCCAACT-CCCTAGATGTGTCGTGGGGGCTGAACGTCGCCGCTTAAAGGGGGGG 317
DB 41349 TCCGCCCAACTCCCTATA-TGTGTCGTGGGGGCTGAACGTCGCCGCTTAAAGGGGGGG 41408

QY 318 CCGCGGTCCACGTCGCTTCTGTCAGTGAAGTAAACATATAACAGAGCGCGGAAGGG 377
DB 41409 CCGCGGTCCACGTCGCTTCTGTCAGTGAAGTAAACATATAACAGAGCGCGGAAGGG 41468

QY 378 GCGGGGAGGAGGAGACACAGGCTTIGACCGATAGTAACCTTCGCTCGGTGCAGCC 437
DB 41469 GCGGGGAGGAGGAGACACAGGCTTIGACCGATAGTAACCTTCGCTCGGTGCAGCC 41528

QY 438 GAATCTATAAAGAACTAGTCCCGGCAAAAACCCCGTAATTCGGAGCGAGTGAAGTGG 497
DB 41529 GAATCTATAAAGAACTAGTCCCGGCAAAAACCCCGTAATTCGGAGCGAGTGAAGTGG 41588

QY 498 GCGCGGACCGCGAGCGGAGCGGACCTTCTCTCCCGGCTCGCGGAGGCGAGCGGG 557
DB 41589 GCGCGGACCGCGAGCGGAGCGGACCTTCTCTCCCGGCTCGCGGAGGCGAGCGGG 41648

QY 558 GGAGCTC 564
DB 41649 GGAGCTC 41655

RESULT 12
LOCUS AX351031 2893 bp DNA linear PAT C6-FEB-2002
DEFINITION Sequence 3 from Patent WO0183746.
ACCESSION AX351031
VERSION AX351031.1 GI:18616387
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P.,
Brewer,B., Duvergier,N., Remaley,A. and Santamarina-Fojo,S.
TITLE Regulatory nucleic acid sequences of the abcl1 gene
JOURNAL Patent: WO 0183746-A 3 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
1..2893
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 745 a 681 c 769 g 598 t
ORIGIN
Query Match 80.9%; Score 456.4; DB 6; Length 2893;
Best Local Similarity 98.7%; Pred. No. 3.9e-99;
Matches 472; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
```


Center project name: L4483
Center clone name: 24_J_9

* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length
* 2825 2924: gap of 100 bp
* 2925 3802: contig of 878 bp in length
* 3803 3902: gap of 100 bp
* 3903 4816: contig of 914 bp in length
* 4817 4916: gap of 100 bp
* 4917 5739: contig of 843 bp in length
* 5760 5859: gap of 100 bp
* 5860 6764: contig of 905 bp in length
* 6765 6864: gap of 100 bp
* 6865 7747: contig of 883 bp in length
* 7748 7847: gap of 100 bp
* 7848 8755: contig of 908 bp in length
* 8756 8855: gap of 100 bp
* 8856 9753: contig of 898 bp in length
* 9754 9853: gap of 100 bp
* 9854 10757: contig of 904 bp in length
* 10758 10857: gap of 100 bp
* 10858 11732: contig of 875 bp in length
* 11733 11832: gap of 100 bp
* 11833 12739: contig of 907 bp in length
* 12740 12839: gap of 100 bp
* 12840 13710: contig of 871 bp in length
* 13711 13810: gap of 100 bp
* 13811 14684: contig of 874 bp in length
* 14685 14784: gap of 100 bp
* 14785 15662: contig of 878 bp in length
* 15663 15762: gap of 100 bp
* 15763 16677: contig of 915 bp in length
* 16678 16777: gap of 100 bp
* 16778 17678: contig of 901 bp in length
* 17679 17778: gap of 100 bp
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18780 19632: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19733 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22573: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23668: gap of 100 bp
* 23669 24554: contig of 866 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length

* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 880 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 887 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp

```
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match          40.7%; Score 229.8; DB 2; Length 90598;
Best Local Similarity 66.0%; Pred. No. 9.7e-45;
Matches 279; Conservative 0; Mismatches 141; Indels 3; Gaps 3;

Qy 142 AGCCGAGCCGCGGACCTTAAGACACCTGCTACCCICACCCCGCCACCCACCTCC 261
Db 87863 AGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87804
Qy 202 AGCGACGCGCGGACCTTAAGACACCTGCTACCCICACCCCGCCACCCACCTCC 261
Db 87803 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87744
Qy 262 CCCCACCTCCCTAGATGTCGTGGCGGCTGAAGCTCGCGCTTAAAGGCGCGGCCCC 321
Db 87743 CCAGTCCCACTCATAGTGTCTGGCGCTGAGGTCGCGC-TTAAGGGCGCGGCC 87895
Qy 322 GGTCCACAGTCTTCTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 381
Db 87684 GGTCCACAGTCTTCTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 87626
Qy 382 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 441
Db 87625 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 87566
Qy 442 CTATAAAGGAACTAGTCCCGGCAAAACCCCGTAATTCCGAGCGAGAGTGAAGTGG 501
Db 87565 CTATAAAGGAACTAGTCCCGGCAAAACCCCGTAATTCCGAGCGAGAGTGAAGTGG 87507
Qy 502 GGGAGCGCGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 561
Db 87506 GGGAGCGCGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 87447
Qy 562 CTC 564
Db 87446 CTC 87444
```

```
RESULT 15
AF258627 AF258627 697 bp mRNA linear PRI 11-MAY-2000
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
DEFINITION partial cds.
ACCESSION AF258627
VERSION AF258627.1 GI:7769707
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 697)
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
Analysis of ABCA1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
2 (bases 1 to 697)
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J. and Kane, J.P.
Direct Submission
Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
```

```
FEATURES
source
Location/Qualifiers
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/contig="ABCA1"
/contig_start=1
/contig_end=697
/feature="membrane-bound"
/product="ATP binding cassette transporter 1"
/protein_id="AA69513.1"
/db_xref="GI:7769707"
/translocation="MACWPOLRLLKMLTFRRTQCOLLELVANPLFLILISVRL
SYPPYEQHECHFPNKAMPAGTLPWVGIIICNANNPCFRYPTGCEAPGVVGNFNS"
BASE COUNT 152 a 198 c 190 g 156 t 1 others
ORIGIN
Query Match 30.7%; Score 173; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 3.8e-31;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 392 AGAGCACAGGCTTTACCGATAGTAACCTCTCGCTCGCTGAGTGGCGCGGACCCCGCA 451
Db 1 AGAGCACAGGCTTTACCGATAGTAACCTCTCGCTCGCTGAGTGGCGCGGACCCCGCA 60
Qy 452 AACTAGTCCCGGCAAAACCCCGTAATTCCGAGCGAGAGTGAAGTGGCGCGGACCCCGCA 511
Db 61 AACTAGTCCCGGCAAAACCCCGTAATTCCGAGCGAGAGTGAAGTGGCGCGGACCCCGCA 120
Qy 512 GAGCGAGCGGACCCCTTCTCTCCCGGGTCTCGCGGAGCGGAGCGGAGCGGAGCTC 564
Db 121 GAGCGAGCGGACCCCTTCTCTCCCGGGTCTCGCGGAGCGGAGCGGAGCGGAGCTC 173
Search completed: February 4, 2003, 01:21:04
Job time : 2000.75 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:39:45 ; Search time 1067.18 Seconds

(without alignments)
8559.210 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1080_1643

Perfect score: 564

Sequence: 1 gagctctctctctccccaac.....cagggcaggcgaggagctc 564

Scoring table: IDENTIFY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmc:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89.4	15.9	736	9	AU135588
2	89	15.8	547	9	AL698654
3	89	15.8	763	9	AU121731
4	75	13.3	292	14	Z44377
5	75	13.3	1004	11	BC034824
6	60.6	10.8	1201	17	CNS015W3

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	60.2	10.7	998	12	BG678861
8	54.8	9.7	298	17	CNS04M2X
9	53.6	9.5	844	17	CNS0052P
c 10	52.4	9.3	910	17	CNS0060N
c 11	51.2	9.1	832	17	CNS015XC
c 12	50.6	9.0	925	17	CNS0091P
c 13	50.6	9.0	1101	17	CNS017SY
c 14	50.4	8.9	870	17	CNS006E2
c 15	49.8	8.8	844	17	CNS0052P
c 16	48.4	8.6	681	17	CNS02FE9
c 17	48.4	8.6	1004	14	BQ944204
c 18	48	8.5	900	17	AG081217
c 19	47.6	8.4	1609	17	CNS010EW
c 20	47.6	8.4	1101	17	CNS01720
c 21	47.4	8.4	1059	14	BQ680479
c 22	47.4	8.4	1101	17	CNS017V2
c 23	47.2	8.4	477	9	AL513813
c 24	47.2	8.4	925	17	CNS0091P
c 25	47.2	8.4	942	17	AG043878
c 26	47	8.3	843	17	CNS00CS1
c 27	46.8	8.3	431	9	AL513947
c 28	46.2	8.2	527	9	AL514325
c 29	46.2	8.2	875	17	AG043475
c 30	46.2	8.2	1136	14	BQ943816
c 31	46.2	8.2	1417	14	BM810236
c 32	46	8.2	758	17	AG073529
c 33	45.8	8.1	523	9	AL514015
c 34	45.8	8.1	766	17	AG041031
c 35	45.8	8.1	932	17	CNS0072Q
c 36	45.8	8.1	997	14	BQ953024
c 37	45.6	8.1	456	9	AL554951
c 38	45.6	8.1	803	17	AG157500
c 39	45.6	8.1	1213	13	BM560246
c 40	45.4	8.0	486	17	CNS00A3H
c 41	45.2	8.0	853	14	BQ219067
c 42	45	8.0	683	12	BG086852
c 43	45	8.0	884	17	CNS018NP
c 44	45	8.0	1139	14	BM914614
c 45	45	8.0	1151	17	CNS024TJ

ALIGNMENTS

RESULT 1	AU135588	736 bp	mRNA	linear	EST 02-AUG-2002
LOCUS	AU135588	PLACEL	Homo sapiens	cdna	clone PLACE1002437 5', mRNA
DEFINITION	AU135588	sequence.			
ACCESSION	AU135588	GI:10996127			
VERSION	AU135588.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 736)				
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isoqai, T.				
TITLE	HRI human cDNA project				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isogai				
	Genomics Laboratory				
	Helix Research Institute				
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan				
	Tel: 81-438-52-3975				
	Fax: 81-438-52-3986				
	Email: genomics@hri.co.jp				
	HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix				
	Research Institute; cDNA library construction: Department of				
	Virology, Institute of Medical Science, University of Tokyo, and				
	Helix Research Institute.				

```

FEATURES
  source      Location/Qualifiers
              1..736
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="PLACE1002437"
              /clone_lib="PLACE1"
              /tissue_type="placenta"
              /note="Vector: pME18SFL3"
BASE COUNT   163 a 199 c 199 g 170 t 5 others
ORIGIN
Query Match   15.9%; Score 89.4; DB 9; Length 736;
Best Local Similarity 98.9%; Pred. No. 1.9e-10;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 474 GTAATTCGACGAGAGTGAGTGGGCGGCGGAGCCGACGAGCGGACCTTCTCTC 533
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GTAGTTGCGACGAGAGTGAGTGGGCGGCGGAGCCGACGAGCGGACCTTCTCTC 60

QY 534 CCGGGCTGCGGCGGCGGCGGCGGAGCTC 564
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CCGGGCTGCGGCGGCGGCGGCGGAGCTC 91

RESULT 2
LOCUS      AL698654 547 bp mRNA linear EST 21-MAR-2002
DEFINITION DKF2686N12109_r1 586 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION AL698654
VERSION AL698654
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686N12109) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
  source      Location/Qualifiers
              1..547
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="DKFZp686N12109"
              /clone_lib="686 (synonym: hlcc3)"
              /tissue_type="human skeletal muscle"
              /dev_stage="adult"
              /lab_host="DH-0B"
              /note="Vector: pT:ip1Ex2; Site_1: Sf1A; Site_2: Sf1B;
              cDNA-collection"
BASE COUNT   112 a 154 c 157 g 124 t
ORIGIN
Query Match   15.8%; Score 89; DB 9; Length 547;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 AATTGCGGACGAGAGTGAGTGGGCGGCGGAGCCGACGAGCGGACCTTCTCTC 535
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

AUTHORS
 Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houlgatte, R., Juneau, M.M., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietsu, G., Pouliot, Y., Sebastiani-Kabatchis, C., and Tessier, A.

TIITLE
 IMAGE: molecular integration of the analysis of the human genome and its expression

JOURNAL
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

MEDLINE
 95277534

COMMENT
 Contact: Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 92002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr

Single read
 Genexpress_library_id: C; Genexpress_sequence_id: ylc-1zb08
 Seq primer: (-21)M13_universal.

FEATURES
 source
 ..292
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-1zb08"
 /clone_lib="normalized infant brain cDNA"
 /sex="female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex:female; dev_stage:3 months old; isolate: muscular atrophy patient; tissue_type: total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B. Soares, Psychiatry Dept., Columbia University, USA. Normalization method: Bento. Soares, P.N.A.S in press"

BASE COUNT
 50 a 87 c 96 g 56 t 3 others

ORIGIN
 Query Match 13.3%; Score 75; DB 14; Length 292;
 Best Local Similarity 100.0%; Pred. No. 4.4e-07;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 GTGAGTGGGGCGGAGCCGAGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCACAGG 549
 |||||
Db 1 GTGAGTGGGGCGGAGCCGAGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCACAGG 60

QY 550 CAGGCGGGGAGCTC 564
 |||||
Db 61 CAGGCGGGGAGCTC 75

RESULT 5
BC034824
LOCUS BC034824 1004 bp mRNA linear HTC 26-JUL-2002
DEFINITION Homo sapiens, clone IMAGE:4749735, mRNA.
ACCESSION BC034824
VERSION BC034824.1 GI:21961568
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1004)
 Strausberg, R.
 Direct Submission
 Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.

cdNA Library Preparation: Life Technologies, Inc.
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Haie, S.M., Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 52 Row: d Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA y1: 5915657
 This clone has the following problem: incomplete processing.

FEATURES
 Location/Qualifiers
 ..1004
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4749735"
 /tissue_type="Skin, squamous cell carcinoma"
 /clone_lib="NCI_CGAP_Skn4"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6-ccdb"

BASE COUNT
 306 a 234 c 230 g 234 t

ORIGIN
 Query Match 13.3%; Score 75; DB 11; Length 1004;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 GTGAGTGGGGCGGAGCCGAGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCACAGG 549
 |||||
Db 1 GTGAGTGGGGCGGAGCCGAGAGCCGAGCCGACCCCTCTCTCCCGGCTCGGCACAGG 60

QY 550 CAGGCGGGGAGCTC 564
 |||||
Db 61 CAGGCGGGGAGCTC 75

RESULT 6
CNS015W3
LOCUS CNS015W3 1201 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC BACN15123 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL105381
VERSION AL105381.1 GI:5619397
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1201)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosAC11.

FEATURES
 Location/Qualifiers
 ..1201
 /organism="Drosophila melanogaster"

```

/db_xref="taxon:7227"
/clone="BACN15123"
/clone_lib="DrosBAC"
/plasmid="pBeoBAC11"
/note="end : SP6"

BASE COUNT      255 a      276 c      278 g      235 t      157 others
ORIGIN
Query Match      10.8%; Score 60.8; DB 17; Length 1201;
Best Local Similarity 37.3%; Pred. No. 0.00055;
Matches 109; Conservative 51; Mismatches 132; Indels 0; Gaps 0;

QY 36 GGAATACTAACAAAGAAAAAATTCGGAAAGCAGGATTAGAGGAAGCAAAATTCAC 95
Db 857 KRAAAMAAAKKTKMTITGKATMKDKTGGMAATAAAGAAAAAAWADDADTTKKK 916
QY 96 TGGTCCCTTGCTCCGGGAACATGGACTAGAGATCTGGCGCGACCCCGAGCCGAG 155
Db 917 TKTGKCGSGSCSGCTSCCGSSCARASMGSGSSGGGCGCCMGCGCCCGSGSCNGC 976
QY 156 CGCTCCCGCGGCTTTAGGCGGGGGCGCCCGGGGGGGAAGGGGACGCGACCGCGGA 215
Db 977 CGSSCCGCCNCCSSSSSGGGGGGGGCGGCGGGGGGGGGGGGGGGGGGGGGGGGG 1036
QY 216 CCCTAAGACACCTGTGTACCTTCACCCACCCACCCACCCACCTCCCCCACTCCCTAG 275
Db 1037 CCCSCCCCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1096
QY 276 ATGTCTGTGGCGCTCAAGCTCCCGCTTAAGGGGGGGGGGGGGGGGGGGGGGGCTCC 327
Db 1097 CCSSSSSSCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1148

RESULT 7
LOCUS BG578861
DEFINITION 602624760f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5',
mRNA sequence.
ACCESSION BG578861
VERSION BG578861.1 GI:13910258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10603 row: 9 column: 16
High quality sequence stop: 860.
FEATURES
Location/Qualifiers
1..998
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749735"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site:1; Not1;
Site:2; SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      285 a      233 c      244 g      236 t

Query Match      9.7%; Score 54.8; DB 17; Length 298;
Best Local Similarity 40.5%; Pred. No. 0.017;
Matches 81; Conservative 33; Mismatches 86; Indels 0; Gaps 0;

QY 101 CCCTTGGTTCGGGACGTGGACTAGAGTCTCGCGCGAGCCCGGAGCCAGCGGT 160
Db 82 CCNKXSCCCVDVGGCCVCSCGSGSGSCSCSCSSSSSCSCSCSCSCSCSCSCSCSC 141
QY 161 CCCGCGGTCTTAGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 220
Db 142 GSGSSCCSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 201

ORIGIN
Query Match      10.7%; Score 60.2; DB 12; Length 998;
Best Local Similarity 89.0%; Pred. No. 0.00078;
Matches 65; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 492 GAGTGGGGCGGGACCGGAGCGGACCGACCTTCTCTCCCGGCTGGGGCAGGCA 551
Db 1 GTGAGTGGGGCGGGACCGGAGCGGACCGGACCGACCTTCTCTCCCGGCTGGGGCAGGCA 60
QY 552 GGGCGGGGAGCTC 564
Db 61 GGGCGGGGAGCTC 73

RESULT 8
LOCUS CNS04M2X
DEFINITION 298 bp DNA linear GSS 21-MAY-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
119N17 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL296898
VERSION AL296898.1 GI:8035478
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 298)
AUTHORS Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 298)
AUTHORS Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 298)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Location/Qualifiers
1..298
Source
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="119N17"
/clone_lib="G"
/note="Genoscope sequence ID : C0BGL19CG09LP1-end : T7"

BASE COUNT      18 a      106 c      80 g      15 t      79 others
ORIGIN
Query Match      9.7%; Score 54.8; DB 17; Length 298;
Best Local Similarity 40.5%; Pred. No. 0.017;
Matches 81; Conservative 33; Mismatches 86; Indels 0; Gaps 0;

QY 101 CCCTTGGTTCGGGACGTGGACTAGAGTCTCGCGCGAGCCCGGAGCCAGCGGT 160
Db 82 CCNKXSCCCVDVGGCCVCSCGSGSGSCSCSCSSSSSCSCSCSCSCSCSCSCSCSC 141
QY 161 CCCGCGGTCTTAGCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 220
Db 142 GSGSSCCSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 201

```



```

QY 221 AGACACCTGCTGCTACCTCCACCCACCCACCCACCCACCTCCCGCAACTCCCTAGATGTG 280
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 NNCNCNCCTCTCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNC 261
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 TCGTGGGCGGCTGAACGTCG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 GGGGGGGGGGGGGGGGGGGG 281

RESULT 9
LOCUS CNS0052P 844 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR11P16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL056652
VERSION AL056652.1 GI:4932342
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 844)
AUTHORS Direct Submission
TITLE Drosophila melanogaster.
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.h.m.

FEATURES
source
location/Qualifiers
1..844
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR11P16"
/clone_lib="RPCI-98"
/note="end : TET3"

BASE COUNT 261 a 112 c 92 g 35 t 344 others
ORIGIN
Query Match 9.5%; Score 53.8; DB 17; Length 844;
Best Local Similarity 23.8%; Pred. No. 0.023;
Matches 86; Conservative 131; Mismatches 140; Indels 5; Gaps 1;

QY 38 AAACATAACAAAGGAAAAAAATTCGGGAAGCAGGATTAGAGCAATTCCTACTG 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 AAAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 98 GTGCCCTTGGCTCGCGGGAAGCTAGAGTCTCGCGCGGCGGCGGCGGCGGCGGCGG 157
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 SASVSSACSSVAGSMAAGACGCGCAANNACGCGSSSSSASRMMSSSCSSCASSSCS 420
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 CTTCCCGCGCTCTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 217
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 VGSSSSASMSGCCCVVSSCGMASSSCGMSASSSSSSSSSSSSSSSSSSSSSSSSSS 480
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 218 CTAAGACACCTGCTGTACCTTCACCCCGCCACCCACCCACCTCCCGCCACTCCCTAGAT 277
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 481 S-----MVCCCMCMCCMCCSVSVCAVCSGSGSVSVAVGCSVGGRRVGSRRGRAGRS 535
QY 278 GTGTCTGGCGGCTGAAGCTCCCGCTTTAAGGGCGGCGGCGGCTCCACCTGCTTTC 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 536 SCRGGSSSVSSGWSSSSVGMGCACSSASVSCBSSSVASSVSGVBVSRGCRVGGGV 595
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 TCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 596 GGGSRVSCSGSSSSSSSSSSSSSGSVSRSSRSANGVYVGGGRRGGGRRGGANARAD 655
QY 398 CA 399
Db 656 DA 657

RESULT 10
LOCUS CNS0050N 910 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION AL065629
VERSION AL065629.1 GI:4944698
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 910)
AUTHORS Direct Submission
TITLE Drosophila melanogaster.
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.h.m.

FEATURES
source
location/Qualifiers
1..910
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR14J21"
/clone_lib="RPCI-98"
/note="end : T7"

BASE COUNT 202 a 63 c 112 g 198 t 335 others
ORIGIN
Query Match 9.3%; Score 52.4; DB 17; Length 910;
Best Local Similarity 20.5%; Pred. No. 0.047;
Matches 48; Conservative 104; Mismatches 82; Indels 0; Gaps 0;

QY 39 AACTACAAAGGAAAAAAATTCGGGAAGCAGGATTAGAGCAATTCCTACTG 98
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 752 AASARVAAVAVVAASAAVAASAAVAASAAVAASAAVAASAAVAASAAVAASAA 693
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 TGCCCTTGGCTCGCGGGAAGCTAGAGTCTCGCGCGGCGGCGGCGGCGGCGGCGG 158
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 692 SASSSASASMSVSCRSAVSAASAAASASASASASASASASASASASASASAS 633
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 159 TTCCCGCGCTCTTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2-8

```



```

RESULT 13
CNS0175Y      CNS0175Y      1101 bp      DNA      linear      GSS 26-JUN-1999
LOCUS         Drosophila melanogaster genome survey sequence SP6 end of SAC
DEFINITION    BAC37L08 of DrosBAC library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION     AL108460
VERSION       AL108460.1 GI:5628764
KEYWORDS      GSS.
SOURCE        Drosophila melanogaster.
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 1101)
AUTHORS       Direct Submission
TITLE         Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT       Determination of this BAC-end sequence was carried out as part of a
              collaboration with the European Drosophila Genome Project (BDGP) -
              http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
              library (Dros BAC) was made by Alain Billaud at CEPH (Centre
              d'Etude du Polymorphisme Humain) with funding provided by a MRC
              project grant. The DNA was prepared from embryos by Alain Bucheton
              and Genevieve Payan. It has been constructed in the vector
              pBelobAC11.
FEATURES       Location/Qualifiers
               1..1101
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="BAC37L08"
                /clone_lib="DrosBAC"
                /plasmid="pBelobAC11"
                /note="end : SP6"
BASE COUNT    254 a 176 c 160 g 152 t 359 others
ORIGIN
Query Match   9.0%; Score 50.6; DB 17; Length 1101;
Best Local Similarity 13.4%; Pred. No. 0.12;
Matches 53; Conservative 187; Mismatches 157; Indels 0; Gaps 0;

Qy 129 GAGTCTGGCGGCGGACGCCGACCGCTTCCGCGCGCTTACGGCGGCGGCGCGG 188
Db 675 GKKAKGCVRGVRCCAGGGGACACMAADCGCCAKMACCCSSSSASSTSSSSAS 734
Qy 189 GCGGGGAAGGGGACGACCGCGGACCTTAAGACACCTCTGTACCTCCACCCGAC 248
Db 735 RGVVSSCACSGSGGASACGAGGGGGRGGGCGCCASGGVCGAACSSASCSMSMC 794
Qy 249 CCCACCCACCTCCCGCAACCTCCCTAGATGTGTGTCGGCGGCTGAAGTCGCCGCTTA 308
Db 795 GVSAGSCSASCSCCGVSSCSAYASASSASVNSKVASAVASCASAVASGAGVSSSCR 854
Qy 309 AGGGCGGGCGCCCGCTCCAGCTCTTCTCTCTGAGTGACTGAATAAAGAGAGG 368
Db 855 SSVASSVSAASVSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 914
Qy 369 CGGAACGGGGCGGGAGGAGGAGACAGAGCTTTGACCGATAGTAACCTCTCCGCTC 428
Db 915 VVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 974
Qy 429 GGTGACCGCGAATCTATAAAGGAACCTAGTCCCGGCAAAAACCCCGTAATTCGCGGAG 488
Db 975 ASSSSSVSVSVVAVAAAAAASAAAAAASSSASAVAVSVASSSSSSSSSSSSS 1034
Qy 489 AGTGGTGGCGGCGGACCGGACGAGCGGACCGGACC 525
Db 1035 SVSSSSSSSSSVSVSSSSSVSVASASASVSSV 1071

RESULT 14

```

```

CNS006EZ/c    CNS006EZ      870 bp      DNA      linear      GSS 03-JUN-1999
LOCUS         Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION    BACR13E15 of RPCI-98 library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION     AL064271
VERSION       AL064271.1 GI:4944346
KEYWORDS      GSS.
SOURCE        Drosophila melanogaster.
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 870)
AUTHORS       Direct Submission
TITLE         Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT       Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazutoyo Osadaawa and
              Aaron Mammoser in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCI-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain y2; cn bw sp, the same strain used for the BDGP's
              p1 and EST libraries. A more detailed description of the library
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
               1..870
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="BACR13E15"
                /clone_lib="RPCI-98"
                /note="end : TET3"
BASE COUNT    256 a 90 c 87 g 127 t 310 others
ORIGIN
Query Match   8.9%; Score 50.4; DB 17; Length 870;
Best Local Similarity 28.0%; Pred. No. 0.13;
Matches 67; Conservative 74; Mismatches 98; Indels 0; Gaps 0;

Qy 95 CIGTGCCCTTGGTGCCTGACCTCCACCCACCCACCCACCCACCCACCTCCCTA 154
Db 809 CSSGTGTTCTSSCTTTCGCKSSGGGTSSSCBTTTCGSCSTCSSSTCTCTTTCCG 750
Qy 155 CGCTTCCCGCGCTCTAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 214
Db 749 GCSSTYSSSTTTCBKSSSGSKANVKSGBSGGCSYVVGSSASMSGCGRASAGKRS 690
Qy 215 ACCCTAACACACCTGCTGTACCTCCACCCACCCACCCACCCACCCACCTCCCTA 274
Db 689 ASSGHGSSCGSSGSGSYSACCCMCSCSCYCCSCSYCCSCCTCTCTTTTGBBESC 630
Qy 275 GATGTGTCGTGGCGGCTGAACGTGCGCCGTTTAAGGGGCGGCGGCGGCGGCGG 333
Db 629 GCTSBKSMGMSGNGTGTCYCCCTCTCCGTCGCGCGCGCGCSCYNCSCCCSCSCSY 571

RESULT 15
CNS0052P      CNS0052P      844 bp      DNA      linear      GSS 03-JUN-1999
LOCUS         Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION    BACR1P16 of RPCI-98 library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION     AL056552
VERSION       AL056552.1 GI:4932342

```

KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 844)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL
COMMENT
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..844
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BAC11P16"
/clone.lib="RPCI-98"
/note="end : REF3"
BASE COUNT 261 a 112 c 92 g 35 t 344 others
ORIGIN
Query Match 8.8%; Score 49.8; DB 17; Length 844;
Best Local Similarity 17.8%; Pred. No. 0.19;
Matches 77; Conservative 173; Mismatches 180; Indels 2; Gaps 1;
QY 1 GAGCTCTCTCCCGCAATCCCTCCCTCGGCTGAGGAACTAACAAAGGAAAAAAT 60
DB 790 DAACCCCGCCCGCCCMCHMCMCMVCCITTTTTCMTYAMMYAMMYAMMYV 731
QY 61 TCGGGAACGAGATTAGAGGAAGCAATTCACGTGTCCTTGCCTCGCGGACGT 120
DB 730 TNCVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMVAMV 671
QY 121 GCACTAGAGAGTCTCGGCGGAGCCCGAGCCCGAGCGCTTCCCGCGCTTTAGCGCGC 180
DB 670 NENTVMTVTYVTHHTYNTCCYCCCGCCYCCBYCBNTSCSYYSYSCSGCSSS 611
QY 181 GGGCCGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 610 SSSCSGSSBYSCCCBCCCBGYSBVCSSBSTBSVSGSSBSTSGTCKCBSSSS 551
QY 241 ACCCCACCCACCCACCCACCCACCTCCCTAGATGTCGTGGCGGCTGAACGTCG 300
DB 550 SBECSSBSCCYCSYCTYCSSC--CBYCCBSCBTSYBSCSSGFTGBSSBSGKSG 493
QY 301 CCGGTTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 360
DB 492 GKSSGKGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 433
QY 361 ACAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 432 SSKSTSGSSBSGSSBSGSSBSGSSBSGSSBSGSSBSGSSBSGSSBSGSSBSGSS 373
QY 421 CTGGGCTCGGTG 432
DB 372 TSSGTSSSSTS 361

Search completed: February 4, 2003, 03:26:21
Job time : 1080.9 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:56:40 ; Search time 26.4714 Seconds
(without alignments)
6534.966 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1080_1643

Perfect score: 364

Sequence: 1 gagctctctctcccccaatc.....caggcgaggcgaggagcgc 564

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38.8	6.9	595	4	US-08-483-533-4
2	38.8	6.9	595	4	US-09-283-471A-4
3	38.8	6.9	1327	4	US-08-483-533-36
4	38.8	6.9	1327	4	US-09-283-471A-36
5	38.8	6.9	1335	5	PCT-US91-06532-1
6	38.6	6.8	4481	4	US-09-041-886-18
7	37.6	6.7	936	1	US-08-018-977C-4
8	36.2	6.4	43280	2	US-08-804-237C-1
9	35.8	6.3	4065	1	US-08-072-574-7
10	35.8	6.3	4181	1	US-08-072-574-9
11	35.8	6.3	11907	4	US-08-061-376-4
12	35.6	6.3	4411523	4	US-09-103-840A-1
13	35.6	6.3	4411523	4	US-09-103-840A-1
14	35.2	6.2	35100	1	US-08-306-691B-19
15	35.2	6.2	35100	5	PCT-US93-06251-19
16	35	6.2	7218	1	US-08-232-463-14
17	34.6	6.1	1280	4	US-08-483-533-38
18	34.6	6.1	1280	4	US-09-283-471A-38
19	34.6	6.1	17410	1	US-07-841-646-3
20	34.6	6.1	17410	1	US-08-147-023-3
21	34.6	6.1	17410	1	US-08-447-570-3
22	34.6	6.1	17410	2	US-08-449-700-3
23	34.6	6.1	17410	2	US-08-449-699A-3
24	34.4	6.1	954	1	US-08-170-294-1
25	34.4	6.1	954	2	US-08-664-855-1
26	34.4	6.1	954	3	US-09-049-289-1
27	34.4	6.1	2888	4	US-08-765-907A-1

C 28	34.2	6.1	1490	4	US-08-964-127-3	Sequence 3, Appli
C 29	34.2	6.1	1490	4	US-09-496-692-3	Sequence 3, Appli
C 30	34.2	6.1	1610	4	US-09-276-531-56	Sequence 56, Appli
C 31	34.2	6.1	2460	4	US-08-964-127-1	Sequence 1, Appli
C 32	34.2	6.1	2460	4	US-09-496-692-1	Sequence 1, Appli
C 33	34.2	6.1	4085	1	US-08-486-270-7	Sequence 7, Appli
C 34	34.2	6.1	4085	3	US-08-367-264-7	Sequence 7, Appli
C 35	34.2	6.1	4085	4	US-09-153-757-7	Sequence 7, Appli
C 36	34.2	6.1	4181	1	US-08-486-270-9	Sequence 9, Appli
C 37	34.2	6.1	4181	3	US-08-367-264-9	Sequence 9, Appli
C 38	34.2	6.1	4181	4	US-09-153-757-9	Sequence 9, Appli
C 39	34.2	6.1	4207	4	US-08-660-148-1	Sequence 1, Appli
C 40	34.2	6.1	4207	4	US-08-660-148-3	Sequence 3, Appli
C 41	34.2	6.1	4303	4	US-08-660-148-4	Sequence 4, Appli
C 42	34.2	6.1	4303	4	US-08-660-148-6	Sequence 6, Appli
C 43	34.2	6.1	72604	4	US-09-268-992-7	Sequence 7, Appli
C 44	34.2	6.1	72604	4	US-09-657-474-7	Sequence 7, Appli
C 45	33.8	6.0	2877	4	US-09-235-103-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-483-533-4
; Sequence 4, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; INVENTOR: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; DISEASES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Brown
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/POCKET NUMBER: 28097/32742
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-483-533-4

Query Match 6.9%; Score 38.8; DB 4; Length 595;

Best Local Similarity 49.28; Score 38.8; DB 4; Length 595;
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
QY 92 CCACGGTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
DB 99 CCGAGAGACACTGGCGGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
QY 152 CCAGGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
DB 159 CCGAGAGACACTGGCGGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
QY 212 CCGAGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 211
DB 219 CCGAGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 211
QY 269 TCCCTAGATGCTGCTGGGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 268
DB 279 TCCCTAGATGCTGCTGGGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 268
QY 329 CGTGTCTTCTGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 354
DB 339 GGGCGCGGAGCGCGGCGGCGT 364

RESULT 2
US-09-283-471A-4
; Sequence 4, Application US/09283471A
; Patent No. 6340673
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,471A
; FILING DATE: 04-APR-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 27373/32742A
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-09-283-471A-4
Query Match 6.94; Score 38.8; DB 4; Length 595;
Best Local Similarity 49.28; Pred. No. 0.083;
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
QY 92 CCACGGTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
DB 99 CCGAGAGACACTGGCGGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 151
QY 152 CCAGGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 211
DB 159 CCGAGAGACACTGGCGGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 211
QY 212 CCGAGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 268
DB 219 CCGAGCTTCGGCGCTTGGCTGCGGGAACGTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 268
QY 269 TCCCTAGATGCTGCTGGGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 328
DB 279 TCCCTAGATGCTGCTGGGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 328
QY 329 CGTGTCTTCTGCTGAGTCTAGAGTCTGGGGCGAGCCCGGAGC 354
DB 339 GGGCGCGGAGCGCGGCGGCGT 364

RESULT 3
US-08-483-533-36
; Sequence 36, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; TITLE OF INVENTION: Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
;
; MOLECULE TYPE: DNA (genomic)
US-08-483-533-36

Query Match
Best Local Similarity 6.9%; Score 38.8; DB 4; Length 1327;
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 92 CCACGTGTCCTGGCTGGCGGAGGTGACGTAGAGAGTCTCGCGCGAGCCCGGAGC 151
D5 540 CCGCAGAGACCTGGCGGCTGCGCTGGCGAGCCGCGGGGAGGGGCGCGGAGC 599
QY 152 CCACGCTTCCCGCGCTCTAGCGCGCGCGCGCGCGGGGGAAGGAGCGAGACCG 211
D5 600 CCGCGGAGACCGCGGACCGCCCGCGAGCCCGCGGAGCCCGCGGACCGCCG 652
QY 212 CGACCGCTAGACACCTGCTGTACCTTCACCTCCACCCACCC---CACCCACCTCCCGGAGC 268
D5 660 CGACCGCGGAGCCCGCGGACCGCGGAGCCCGCGGAGCCCGCGGAGCCCGG 719
QY 269 TCCCTAGATGTGTGTCGGGCGCTGAACGTCGCCGCTTAAAGGGGCGCGCGGCTCCA 328
D5 720 TCCGGGTGCGGACCTGGTGTCTGGGCTTGGCGCGCGCGCGCGCGGCTCGT 779
QY 329 CGTCTTTCTGCTGAGTGACTGAACCT 354
D5 780 GGGCCCGGAGCGGCGGCGGCGCT 805

RESULT 4
US-09-283-471A-36
; Sequence 36, Application US/09283471A
; Patent No. 6340673
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chol, Joany
; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,471A
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 27373/32742A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 36:
; MOLECULE TYPE: DNA (genomic)
; LENGTH: 1327 base pairs
```

```
;
; TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-283-471A-36

Query Match
Best Local Similarity 6.9%; Score 38.8; DB 4; Length 1327;
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 92 CCACGTGTCCTGGCTGGCGGAGGTGACGTAGAGAGTCTCGCGCGAGCCCGGAGC 151
D5 540 CCGCAGAGACCTGGCGGCTGCGCTGGCGAGCCGCGGGGAGGGGCGCGGAGC 599
QY 152 CCACGCTTCCCGCGCTCTAGCGCGCGCGCGCGGGGGAAGGAGCGAGACCG 211
D5 600 CCGCGGAGACCGCGGACCGCCCGCGAGCCCGCGGAGCCCGCGGAGCCCGG 659
QY 212 CGGACCGCTAAGACACCTGCTGTACCTTCCACCCACCC---CACCCACCTCCCGGAGC 268
D5 660 CGACCGCGGAGCCCGCGGAGCCCGCGGAGCCCGCGGAGCCCGCGGAGCCCGG 719
QY 269 TCCCTAGATGTGTGTCGGGCGCTGAACGTCGCCGCTTAAAGGGGCGCGGCTCCA 328
D5 720 TCCGGGTGCGGACCTGGTGTCTGGGCTTGGCGCGCGCGCGCGGCTCGT 779
QY 329 CGTCTTTCTGCTGAGTGACTGAACCT 354
D5 780 GGGCCCGGAGCGGCGGCGGCGCT 805

RESULT 5
PCT-US91-06532-1
; Sequence 1, Application PC/TUS9106532
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza Suite 2100
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06532
; FILING DATE: 19910910
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27373/8235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US91-06532-1
```



```
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS JR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-19

Query Match      6.28; Score 35.2; DB 5; Length 35100;
Best Local Similarity 50.08; Pred. No. 4;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 158 CTTCCCGCGGCTCTTAGCGCGGCGCGCGGCGGGAAGGAGACGACGACCGCGGACC 217
      |||||
Db 8471 CCTGCCCGGACACCTTGTCAGTCCTGCCCCCATAGATGATCGTGGACACCATTTGAGGGGAAG 8412
      |||||

Qy 218 CTAAAGACACCTGTGTACCCCTCCACCCCGGACCCACCGACCTCCCCCGAACTCCCTAGAT 277
      |||||
Db 8411 AAGTGAAAGAAAGAGAGACATCCCACTGCCCGCCCATCACACCCAGTCCCGCAGTCACCCACC 8352
      |||||

Qy 278 GTGTGTTGGGGGCTGTGAACGTCGCCCGTTTAAAGGGGGGGCGCGGCTCCACGTGC 333
      |||||
Db 8351 ATGTGGCGCGCGGTCTCTGTGTCGCCACTCTTCAGGCCGTGGCTGCGGGTCCCATGC 8296
      |||||
```

Search completed: February 3, 2003, 18:51:27
Job time : 8913.47 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 ; Search time 1325.59 Seconds
(without alignments)
10165.006 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1181_1643

Perfect score: 463

Sequence: 1 ccttgctgcgcgggaactgtg.....caggggcagggcgaggagctc 463

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmul: *

1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_cv: *
23: em_pat: *
24: em_ph: *
25: em_pi: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pin: *
35: em_htg_rnd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htg_hum: *
40: em_htg_mus: *
41: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	463	100.0	1167	9	HS252201	AJ252201 Homo sapi
2	463	100.0	1643	6	AX060715	AX060715 Sequence
3	463	100.0	1643	6	AX060894	AX060894 Sequence
c	463	100.0	96717	9	AL359182	AL359182 Human DNA
5	463	100.0	175064	2	AC012230	AC012230 Homo sapi
6	459.8	99.3	1167	9	AF258623S1	AF258623 Homo sapi
7	448.8	96.9	69570	2	AC021246	AC021246 Homo sapi
8	447	96.5	183999	6	AX025589	AX025589 Sequence
9	446.4	96.4	3231	6	AX351029	AX351029 Sequence
10	446.4	96.4	149034	9	AF275948	AF275948 Homo sapi
11	441.6	95.4	201144	9	AF287262	AF287262 Homo sapi
12	355.4	76.8	2893	6	AX351031	AX351031 Sequence
c	271.6	58.7	69570	2	AC021246	AC021246 Homo sapi
c	229.8	49.6	90898	2	AC021345	AC021345 Homo sapi
15	173	37.4	857	9	AF258627	AF258627 Homo sapi
c	155.4	33.6	186889	2	AL807243	AL807243 Mus muscu
17	155.4	33.6	278572	10	AF287263	AF287263 Mus muscu
c	153.2	33.1	145833	2	AC125837	AC125837 Rattus no
19	98	21.2	7260	6	AX253452	AX253452 Sequence
20	91	19.7	221	6	AX351032	AX351032 Sequence
21	89.4	19.3	1556	9	AK024328	AK024328 Homo sapi
22	89	19.2	1750	9	AK022254	AK022254 Homo sapi
23	75	16.2	9854	6	AX127831	AX127831 Sequence
24	75	16.2	9854	6	AX139818	AX139818 Sequence
25	67	14.5	10442	6	AX060713	AX060713 Sequence
26	67	14.5	10442	6	AX060892	AX060892 Sequence
27	67	14.5	10442	9	AF285167	AF285167 Homo sapi
28	60	13.0	10474	6	AX060719	AX060719 Sequence
29	60	13.0	10474	6	AX060721	AX060721 Sequence
30	60	13.0	10474	6	AX060898	AX060898 Sequence
31	60	13.0	10474	6	AX060900	AX060900 Sequence
c	49.6	10.7	96136	2	AC097674	AC097674 Rattus no
c	49.4	10.7	173127	2	AC131142	AC131142 Rattus no
34	49	10.6	65780	2	AC111741	AC111741 Rattus no
35	49	10.6	125020	9	AF429315	AF429315 Homo sapi
c	47.4	10.2	180461	2	AC106563	AC106563 Rattus no
37	47	10.2	167886	2	AC126316	AC126316 Rattus no
c	46.6	10.1	55357	2	AC121553	AC121553 Mus muscu
c	46.4	10.0	184482	2	AC127041	AC127041 Rattus no
c	46.2	10.0	150305	2	AC114705	AC114705 Rattus no
c	46.2	10.0	209216	2	AC117126	AC117126 Rattus no
42	46	9.9	134021	2	AC109884	AC109884 Rattus no
c	46	9.9	174096	2	AC097177	AC097177 Rattus no
c	45.6	9.8	110000	2	AC096324_0	AC096324 Rattus no
c	45.6	9.8	110000	2	AC096324_1	Continuation (2 of

ALIGNMENTS

RESULT 1
HSA252201
LOCUS
DEFINITION
Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, 5'UTR and promoter region.
ACCESSION
AJ252201
KEYWORDS
ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1167)
Porsch-Ozcurumez, M., Langmann, T., Heimerl, S., Borsukova, H.,
Kaminski, W.E., Drobnik, W., Honer, C., Schumacher, C. and Schmitz, G.


```

Db 17329 CTTTGACCATAGTAACTCTCGCTCGCTGGTGGAGCCGCAATCTATAAAGAACIACTGCC 17270
QY 361 GGCAGAAACCCCGTAAATTCGCGAGCGAGAGTAGTGGGGCGGAGCCGCGAGCCGAGCC 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17269 GGCAGAAACCCCGTAAATTCGCGAGCGAGAGTAGTGGGGCGGAGCCGCGAGCCGAGCC 17210
QY 421 GACCCCTCTCTCCCGGGCTGGCGCAGGCGAGGCGGGGAGCTC 463
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17209 GACCCCTCTCTCCCGGGCTGGCGCAGGCGAGGCGGGGAGCTC 17167

RESULT 5
AC012230
LOCUS
DEFINITION
Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
pieces.
AC012230
VERSION
AC012230.3 GI:7637254
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1M10
Unpublished
2 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gace,D., Horton,L.,
Gallagan,J., Gardyna,S., Grant,S., Hages,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatsis,A., Klein,J.,
Lehoczky,J., Lieuw,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,K., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo.A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RN/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2510
Center clone name: 1.M.10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator 319 Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-contents
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contents
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
 * this record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1003: contig of 1003 bp in length
 1004 1103: gap of 100 bp
 1104 2634: contig of 1531 bp in length
 2635 2734: gap of 100 bp
 2735 4415: contig of 1681 bp in length
 4416 4515: gap of 100 bp
 4516 5785: contig of 1270 bp in length
 5786 5885: gap of 100 bp
 5886 7879: contig of 1994 bp in length
 7880 7979: gap of 100 bp
 7980 9686: contig of 1707 bp in length
 9687 9786: gap of 100 bp
 9787 12353: contig of 2467 bp in length
 12354 12353: gap of 100 bp
 12354 15228: contig of 2875 bp in length
 15229 15328: gap of 100 bp
 15329 17200: contig of 1872 bp in length
 17201 17300: gap of 100 bp
 17301 20131: contig of 2831 bp in length
 20132 20231: gap of 100 bp
 20232 22587: contig of 2356 bp in length
 22588 22687: gap of 100 bp
 22688 25707: contig of 3020 bp in length
 25708 25807: gap of 100 bp
 25808 28184: contig of 2377 bp in length
 28185 28284: gap of 100 bp
 28285 31338: contig of 3054 bp in length
 31339 31438: gap of 100 bp
 31439 34299: contig of 2861 bp in length
 34300 34399: gap of 100 bp
 34400 38318: contig of 3919 bp in length
 38319 38418: gap of 100 bp
 38419 42835: contig of 4417 bp in length
 42836 42935: gap of 100 bp
 42936 45448: contig of 2513 bp in length
 45449 45548: gap of 100 bp
 45549 48116: contig of 2568 bp in length
 48117 48216: gap of 100 bp
 48217 52618: contig of 4402 bp in length
 52619 52718: gap of 100 bp
 52719 56592: contig of 3874 bp in length
 56593 56692: gap of 100 bp
 56693 59635: contig of 2843 bp in length
 59636 59735: gap of 100 bp
 59736 63661: contig of 3926 bp in length
 63662 63761: gap of 100 bp
 63762 68437: contig of 4676 bp in length
 68438 68537: gap of 100 bp
 68538 71458: contig of 2921 bp in length
 71459 71558: gap of 100 bp
 71559 76888: contig of 5330 bp in length
 76889 76988: gap of 100 bp
 76989 82113: contig of 5125 bp in length
 82114 82213: gap of 100 bp
 82214 88220: contig of 6007 bp in length
 88221 88320: gap of 100 bp
 88321 93499: contig of 5179 bp in length
 93500 93599: gap of 100 bp
 93600 97901: contig of 4302 bp in length
 97902 98001: gap of 100 bp
 98002 103016: contig of 5015 bp in length
 103017 103116: gap of 100 bp
 103117 109178: contig of 6062 bp in length
 109179 109278: gap of 100 bp
 109279 117307: contig of 8029 bp in length
 117308 117407: gap of 100 bp
 117408 124079: contig of 6672 bp in length
 124080 124179: gap of 100 bp
 124180 131281: contig of 7102 bp in length

* 131282 131381: gap of 100 bp
 * 131382 138059: contig of 6678 bp in length
 * 138060 138159: gap of 100 bp
 * 138160 145491: contig of 7332 bp in length
 * 145492 145591: gap of 100 bp
 * 145592 157491: contig of 11800 bp in length
 * 157392 157491: gap of 100 bp
 * 157492 175064: contig of 17573 bp in length.

FEATURES

Source
 1. 175064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RP11-IM10"
 /clone_lib="RP11-1: Human Male BAC"
 1. 1003
 /note="assembly_fragment"
 1104. 2634
 /note="assembly_fragment"
 2735. 4415
 /note="assembly_fragment"
 4516. 5785
 /note="assembly_fragment"
 5386. 7879
 /note="assembly_fragment"
 7980. 9686
 /note="assembly_fragment"
 9787. 12253
 /note="assembly_fragment"
 12354. 15228
 /note="assembly_fragment"
 15329. 17200
 /note="assembly_fragment"
 17301. 20131
 /note="assembly_fragment"
 20232. 22587
 /note="assembly_fragment"
 22688. 25707
 /note="assembly_fragment"
 25808. 28184
 /note="assembly_fragment"
 28285. 31338
 /note="assembly_fragment"
 31439. 34299
 /note="assembly_fragment"
 34400. 38318
 /note="assembly_fragment"
 38419. 42835
 /note="assembly_fragment"
 42936. 45448
 /note="assembly_fragment"
 45549. 48116
 /note="assembly_fragment"
 48217. 52618
 /note="assembly_fragment"
 52719. 56592
 /note="assembly_fragment"
 56693. 59635
 /note="assembly_fragment"
 59736. 63661
 /note="assembly_fragment"
 63762. 68437
 /note="assembly_fragment"
 68538. 71458
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left"
 71559. 76888

Query Match 100.0%; Score 463; DB 2; Length 175064;
 Best local similarity 100.0%; Pred. No. 2.8e-94;
 Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTGGCTGCCGGGAAGCTGAGTAGAGACTCTCGCGGAGCCGCCAGCCAGCCGCTTC 60

Db	3290	CCTTGGCTGCCGGGAAGCTGAGTAGAGACTCTCGCGGAGCCGCCAGCCGCTTC	3349
QY	61	CCGCGCGTCTTAGCGCGGGGGCGGGGGAGAGGACGACGCGGACCTAA	120
Db	3350	CCGCGCGTCTTAGCGCGGGGGCGGGGGAGAGGACGACGCGGACCTAA	3409
QY	121	GACACCTGTGTACCTCCACCCGCCACCCACACCTCCCTAGATGTG	180
Db	3410	GACACCTGTGTACCTCCACCCGCCACCCACACCTCCCTAGATGTG	3469
QY	181	CGTGGCGGCTGAACGTGCGCCGTTTAAAGGGCGGGCGCGGCTCCACGNGCTTCTGCT	240
Db	3470	CGTGGCGGCTGAACGTGCGCCGTTTAAAGGGCGGGCGCGGCTCCACGNGCTTCTGCT	3529
QY	241	GAGTGACTGAACCTACATTAACAGAGCGCGGGAACGGGGGGAGGAGGAGACGACAGG	300
Db	3530	GAGTGACTGAACCTACATTAACAGAGCGCGGGAACGGGGGGAGGAGGAGACGACAGG	3589
QY	301	CTTTGACCGATAGTAACCTCTGCGCTCGCGAGCGCAATCTATAAAGGAACCTAGTCCC	360
Db	3590	CTTTGACCGATAGTAACCTCTGCGCTCGCGAGCGCAATCTATAAAGGAACCTAGTCCC	3649
QY	361	GGCAAAACCCCGTAATTCGGAGGAGAGTGTGGGCGGGGACCCCGAGACCGAGCC	420
Db	3650	GGCAAAACCCCGTAATTCGGAGGAGAGTGTGGGCGGGGACCCCGAGACCGAGCC	3709
QY	421	GACCTCTCTCCCGGCTCGCGAGGAGCGGCGGAGGCGGAGGCTC	463
Db	3710	GACCTCTCTCCCGGCTCGCGAGGAGCGGCGGAGGAGGCTC	3752

AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
 Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
 promoter and exon 1.
 AF258623
 AF258623.2 GI:8677405
 1 of 4
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1167)
 Roulierat,B.E., Hakamata,H., Duchateau,P.N., Eng,C.,
 Analysis of hABC1 gene 5' end: additional peptide sequence,
 promoter region, and four polymorphisms
 Biochem. Biophys. Res. Commun. 271 (2000) In press
 2 (bases 224 to 1167)
 Roulierat,B.E., Hakamata,H., Duchateau,P.N., Eng,C.,
 Direct Submission
 Submitted (20-APR-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 3 (bases 1 to 1167)
 Roulierat,B.E., Hakamata,H., Duchateau,P.N., Eng,C.,
 Direct Submission
 Submitted (23-JUN-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 Sequence update by submitter
 On Jun 23, 2000 this sequence version replaced gi:7769713.
 REMARK
 COMMENT
 FEATURES
 source
 1. 1167
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"

```

promoter      224..844
/ gene="ABCA1"
exon          845..1147
/ gene="ABCA1"
/ number=1

BASE COUNT   278 a 313 c 328 g 248 t
ORIGIN
Query Match   99.3%; Score 459.8; DB 9; Length 1167;
Best Local Similarity 99.5%; Pred. No. 2.5e-93;
Matches 461; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQTGGCTGCGGGAAGCTGGAGTCTGCGGCGAGCCCGAGCCGCGCTTC 60
Db 555 CTTGGCTGCGGGAAGCTGGAGTCTGCGGCGAGCCCGAGCCGCGCTTC 614
QY 61 CCGCGGCTCTTAGCGCGGCGGCGGCGGCGGGAAGGAGCGAGCGCGACCTAA 120
Db 615 CCGCGGCTCTTAGCGCGGCGGCGGCGGCGGGAAGGAGCGAGCGCGACCTAA 674
QY 121 GACACCTGCTGTACCTCCACCCACCCACCCACCTCCCTCCCACTCCCTAGATGCT 180
Db 675 GACACCTGCTGTACCTCCACCCACCCACCCACCTCCCTCCCACTCCCTAGATGCT 734
QY 181 CGTGGCGGCTGTAACGTCGCGCGCTTTAAGGGCGGCGGCGGCTCCACGTCGTTCTGCT 240
Db 735 CGTGGCGGCTGTAACGTCGCGCGCTTTAAGGGCGGCGGCGGCTCCACGTCGTTCTGCT 794
QY 241 GAGTGACTGACTACATAACACAGAGCGCGGAGCGGCGGCGGAGGAGGAGACACAGG 300
Db 795 GAGTGACTGACTACATAACACAGAGCGCGGAGCGGCGGCGGAGGAGGAGACACAGG 854
QY 301 CTTTGACCGATAGTAACCTCTGCGCTCGGTCGACGCCGAATCTATAAAGGAACCTAGTCCC 360
Db 855 CTTTGACCGATAGTAACCTCTGCGCTCGGTCGACGCCGAATCTATAAAGGAACCTAGTCCC 914
QY 361 GCGAAAACCCCGCTAATTGGGAGCGAGTGTAGTGGGGCGGGAGCCCGCAGCGCGAGCC 420
Db 915 GCGAAAACCCCGCTAATTGGGAGCGAGTGTAGTGGGGCGGGAGCCCGCAGCGCGAGCC 974
QY 421 GACCTTCTCTCCCGGGCTGGCGGAGGCGGCGGCGGAGGCTC 463
Db 975 GACCTTCTCTCCCGGGCTGGCGGAGGCGGCGGCGGAGGCTC 1017

RESULT 7
AC021246      59570 bp   DNA   linear   HTG 13-JUL-2000
LOCUS        Homo sapiens clone RP11-IN10, LOW-PASS SEQUENCE SAMPLING.
AC021246
ACCESSION
VERSION      AC021246.2 GI:9119882
KEYWORDS     HTG; HTGS; PHASE0.
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abramson,H., Allen,N.,
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      1 (bases 1 to 69570)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens Chromosome, Clone RP11-IN10
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 69570)
TITLE        Birren,B., Linton,L., Nusbaum,C., Lander,E., Abramson,H., Allen,N.,
AUTHORS      Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
TITLE        Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
AUTHORS      Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
TITLE        Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
AUTHORS      Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
TITLE        Garcina,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
AUTHORS      Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
TITLE        Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
AUTHORS      Macdonald,P., Marquis,N., McEwan,P., McEwan,P., McKernan,K.,
TITLE        McPheters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
AUTHORS      Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,

```

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Vassiliev,E., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seg.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L2512
Center clone name: L_N_10

* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

1 871: contig of 871 bp in length
* 872 871: gap of 100 bp
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length
* 3746 3845: gap of 100 bp
* 3846 4696: contig of 851 bp in length
* 4697 4796: gap of 100 bp
* 4797 5640: contig of 844 bp in length
* 5641 5740: gap of 100 bp
* 5741 6540: contig of 800 bp in length
* 6541 6640: gap of 100 bp
* 6641 7509: contig of 869 bp in length
* 7510 7609: gap of 100 bp
* 7610 8479: contig of 870 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 880 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 15123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 866 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length

```

*	1967	20066:	gap of	100 bp
*	2067	20291:	contig of 855 bp	in length
*	20922	21021:	gap of	100 bp
*	21022	21865:	contig of 844 bp	in length
*	21866	21965:	gap of	100 bp
*	21966	22932:	contig of 867 bp	in length
*	22933	22932:	gap of	100 bp
*	22933	23780:	contig of 845 bp	in length
*	23781	23880:	gap of	100 bp
*	23881	24733:	contig of 853 bp	in length
*	24734	24833:	gap of	100 bp
*	24834	25670:	contig of 837 bp	in length
*	25671	25770:	gap of	100 bp
*	25771	26621:	contig of 851 bp	in length
*	26622	26721:	gap of	100 bp
*	26722	27576:	contig of 855 bp	in length
*	27577	27676:	gap of	100 bp
*	27677	28532:	contig of 856 bp	in length
*	28533	28632:	gap of	100 bp
*	28633	29492:	contig of 860 bp	in length
*	29493	29592:	gap of	100 bp
*	29593	30455:	contig of 863 bp	in length
*	30456	30555:	gap of	100 bp
*	30556	31410:	contig of 855 bp	in length
*	31411	31510:	gap of	100 bp
*	31511	32368:	contig of 858 bp	in length
*	32369	32468:	gap of	100 bp
*	32469	33312:	contig of 844 bp	in length
*	33313	33412:	gap of	100 bp
*	33413	34268:	contig of 856 bp	in length
*	34269	34368:	gap of	100 bp
*	34369	35204:	contig of 836 bp	in length
*	35205	35304:	gap of	100 bp
*	35305	36156:	contig of 852 bp	in length
*	36157	36256:	gap of	100 bp
*	36257	37128:	contig of 872 bp	in length
*	37129	37228:	gap of	100 bp
*	37229	38093:	contig of 855 bp	in length
*	38094	38183:	gap of	100 bp
*	38184	39031:	contig of 848 bp	in length
*	39032	39131:	gap of	100 bp
*	39132	40006:	contig of 875 bp	in length
*	40007	40106:	gap of	100 bp
*	40107	40967:	contig of 861 bp	in length
*	40968	41067:	gap of	100 bp
*	41068	41913:	contig of 846 bp	in length
*	41914	42013:	gap of	100 bp
*	42014	42824:	contig of 811 bp	in length
*	42825	42924:	gap of	100 bp
*	42925	43776:	contig of 852 bp	in length
*	43777	43976:	gap of	100 bp
*	43978	44752:	contig of 876 bp	in length
*	44753	44952:	gap of	100 bp
*	44953	45724:	contig of 872 bp	in length
*	45725	45824:	gap of	100 bp
*	45825	46643:	contig of 819 bp	in length
*	46644	47343:	gap of	100 bp
*	47344	47599:	contig of 856 bp	in length
*	47599	48551:	contig of 852 bp	in length
*	48552	48651:	gap of	100 bp
*	48652	49485:	contig of 834 bp	in length
*	49486	49585:	gap of	100 bp
*	49586	50440:	contig of 855 bp	in length
*	50441	50940:	gap of	100 bp
*	50941	51404:	contig of 864 bp	in length
*	51405	51504:	gap of	100 bp
*	51505	52372:	contig of 868 bp	in length
*	52373	52472:	gap of	100 bp
*	52473	53328:	contig of 856 bp	in length
*	53329	53428:	gap of	100 bp
*	53429	54268:	contig of 840 bp	in length
*	54269	54368:	gap of	100 bp

Query Match 96.98; Score 448.8; DB 2; Length 69570;

Query Match 50.5%; SCOP 490.6%; MS 2; Mengesha 0.0076;
Best Local Similarity 29.4%; Pred. No. 4.8e-91;
Matches 461; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy	1	CCCTGGCTGCCGGGACGTGGACTATAGAGAGTCTCGCGCGAGCCCGAGCCGAGCGCTTC	60
Db	41192	CCCTGGCTGCCGGGAACTGGACTATAGAGTCTCGGCGAGCCCGAGCCGAGCGCTTC	41251
Qy	61	CCGCGGCTCTTAGCGCGGCGGCCCGGGCGGGGAAGGGGACGACGACCGGACCGCTAA	120
Db	41252	CCGCGGCTCTTAGCCGCGGCGGCCCGGGCGGGGAGGGGACGACGACCGGACCGCTAA	41311
Qy	121	GACACCTGCTGTACCTCCOACCCCGACCCACCCACCTCCGCCCACT-CCCTAGATGTG	179
Db	41312	GACACCTGCTGTACCTCCOACCCCGACCCACCCACCTCCGCCCACTCCCTATATGTG	41371
Qy	180	TCCTGGGCGGCTGAAGCTGGCCGGCTTCTAGGGGGGGGCCCGCGCTCCACGCTTTCTGC	239
Db	41372	TCCTGGGCGGCTGAACGCTGCCCGCTTCTAGGGGGGGGCCCGCGCTCCACGCTTTCTGC	41431
Qy	240	TGAGTGACTGAACCTACATAAACAAGAGCGCGGGAACGGGGCGGGAGGAGAGACACAG	299
Db	41432	TGAGTGACTGAACCTACATAAACAAGAGCGCGGGAACGGGGCGGGAGGAGAGACACAG	41491
Qy	300	GCCTTGACCGCATAGTAACCTCTGGCGTGGTGCACCGCAATCTATAAAGGAAGTACTCC	359
Db	41492	GCCTTGACCGCATAGTAACCTCTGGCGTGGTGCACCGCAATCTATAAAGGAAGTACTCC	41551
Qy	360	CGGCAAAAACCCCGTAATTCCGAGGCGAGAGTGAATGGGGCGCGGACCCCGAGAGCCGACG	419
Db	41552	CGGCAAAAACCCCGTAATTCCGAGGCGAGAGTGAATGGGGCGCGGACCCCGAGAGCCGACG	41611
Qy	420	CGACCTTTCTCTCCCGGGTGTGGGAGGCGAGGGCGGGAGCTC	463
Db	41612	CGACCTTTCTCTCCCGGGTGTGGGAGGCGAGGGCGGGAGCTC	41655

RESULT 8

AX092589

LOCUS

LOCUS
DEFINITE

DEFINITION
ACCESS TO

ACCESSION
NUMBER

VERSION
REVISED

KEYWORDS

100

AX092589	183999 bp	DNA	linear	PAT 21-MAR-2001
Sequence 1 from Patent WO0015676.				

SEQUENCE LISTING

AX092589
22002500
C F - 211111

AX092589.1 GI:13444647


```

source      1. .149034
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
gene         1454. .148034
             /gene="ABCA1"
mRNA         join(1454. .1674,25831. .25959,40385. .40478,45012. .45153,
             45423. .46541,67729. .67850,70831. .71007,83960. .84052,
             89010. .89250,91952. .92101,92433. .92549,96758. .96955,
             97702. .97907,98428. .98604,100391. .100613,102360. .102581,
             10362. .103846,104951. .105064,106862. .107033,
             108023. .108154,109460. .109602,109806. .109943,
             110346. .110866,112124. .112196,113183. .113385,
             115053. .115101,115297. .115410,116806. .116954,
             118604. .118728,119964. .120062,123094. .123283,
             124804. .124898,126208. .126240,127363. .127468,
             128943. .129017,129539. .129708,130939. .131116,
             133114. .133229,133341. .133485,134526. .134649,
             135736. .135865,136130. .136250,137038. .137100,
             138008. .138114,140469. .140510,140982. .141116,
             142060. .142163,142646. .142738,143397. .143640,
             144581. .148034)
             /gene="ABCA1"
             /product="ABCA1"
             join(1454. .1674,25831. .25923)
5'UTR       /gene="ABCA1"
             2406. .2678
repeat_region /rpt_family="Alu"
repeat_region 4240. .4509
             /rpt_family="Alu"
repeat_region 4998. .5270
             /rpt_family="Alu"
repeat_region 5563. .5839
             /rpt_family="Alu"
repeat_region 6615. .6877
             /rpt_family="Alu"
repeat_region 8600. .9046
             /rpt_family="Alu"
repeat_region 9998. .10279
             /rpt_family="Alu"
repeat_region 11859. .12133
             /rpt_family="Alu"
repeat_region 12810. .12902
             /rpt_family="Alu"
repeat_region 15220. .15399
             /rpt_family="Alu"
repeat_region 15508. .15787
             /rpt_family="Alu"
repeat_region 18600. .18860
             /rpt_family="Alu"
repeat_region 20950. .21206
             /rpt_family="Alu"
repeat_region 25038. .25314
             /rpt_family="Alu"
CDS         join(25924. .25989,40385. .40478,45012. .45153,46423. .46541,
             67729. .67850,70831. .71007,83960. .84052,89010. .89250,
             91962. .92101,92433. .92549,96758. .96955,97702. .97907,
             98428. .98604,100391. .100613,102360. .102581,103642. .103846,
             104951. .105064,106862. .107033,108023. .108154,
             109460. .109602,109806. .109943,110646. .110866,
             112124. .112196,113183. .113385,115053. .115101,
             115297. .115410,116806. .116954,118604. .118728,
             119964. .120062,123094. .123283,124804. .124898,
             126208. .126240,127363. .127468,128943. .129017,
             129539. .129708,130939. .131116,133114. .133229,
             133341. .133485,134526. .134649,135736. .135865,
             136130. .136250,137038. .137100,138008. .138114,
             140469. .140610,140982. .141116,142060. .142163,
             142646. .142738,143397. .143640,144581. .144721)
             /gene="ABCA1"
             /codon_start=1
             /product="ABCA1"
             /protein_id="AA886276.1"
             /db_xref="GI:9247086"
/translation="MACPQQLRLKLNLTPTRRROTQQLLEVAWPEFTFLILISVRL
SYPPEQHECEHPNKAMPAGTLPWQGIICNANNEPCREYPTPGAPGVGNKSTIV
ARLFDARLLIYSKQDTSKMDRKVRVLTLOIQIKSSSNLKLQDFLQVNETSGFLYH
KLSLPKSTVDKMRADVILHKVFLQGYQLHLTSLONGSKSEEMIQGDQEVSEDCGLP
REKLAARFVLSNMDILKPIRLTNSPTSPFKELAEATKTLHSLGTLAGELFSMR
SWSMDRQEVMTLVNNSSSSTQIYCAVSRIVCIGHEGGGLXIKSLNMYEDNNYKALF
GNGTEADAETFTYDNSTTTPYCNLAKNLESPLSRIWKALKPLLVGKILYPTDTPAT
ROMVAENKTFQELAVFEDLEGMEWELSPKINTFMENSCQMDLVRMLLDSRDNDFWE
QQLDGLDWAQDIAFLAKHPEDVSSNGSVYTREANETINCAITISRNECVNLN
KLEPIATEVWLINKMELLDERKFWAGVFTGTPGSELPHPHYKIRKMDINDVERT
NLIKDGWDYDPADPFDMDRYWGGFAYLQDVQDAITRVLTGTEKTGYVMQMPY
PCYVDITFLVMSRMPLEMTLAWIYSVAIIKGIYVEKRLKRTMTIMEGDNLSILW
FSWFISLPIVLVSAGLLVILKGNLLPYSDPVSVFVFLSVFAVVTILQCLISLIF
SRANLAAOGGIIYFTLYLVYLCVAVWDYVGFTLKIAXLLSPVAFGFGCYGALFE
EOGLGVQMDNLPESVEEDGFNLTTSVSMLEDTFLYGVMIWYIEAVPGCYGIPRPW
YFPCISYWFGEESDEKSPGNQARISICMEEPPIHLKGVSIONLVKVRDGNKV
AVDGLALNFYEGGITSFLGNGAGKTTMSLTITGLFFPTSGTAYILGKDIRSEMTIR
QNLGVCQHNVLFDMLTVEEHWFYARLKGLEKEVKAEQMDALDVLPSKLSKT
SOLSGMQRLKVALAFVGGSKVILDEPTAGVDVPSRRGIWELLKYRQGTIILST
HMDRADYLGDRITATSHGKLCVGGSLFKNLQGTGYLTLVKKDVSSLSLSSRNS
STVSYLKXEDSVSSSDAGLSDHESPTLTDVSAISNLIKHYSEARLYEDIGSEL
TYVLPYEAKEGAFVELFHEIDRLSDLGTSYGSSETLLEIFUKVAEESVDRET
DGTLPARNRRAGFDKQSCRLPFTEDDAADNDSIDIPRETDLLSGDKGKGYOVK
GWLKTCQFVALLWKRLZARRSRKGFQIYLPVAFVCIALVFLIYFPFGKPSLE
LQPMYNEQYTFVSDNADPDTGTLELLNALTQDPGFGTRCMEGNTIPDTPOAGREW
TTAPVPTIMDLFQNGWNTWNPSPACQSSDKIKKMLPVCPPGAGGLPPQPKONTA
DILODLTGNIISDYLKTYVQIIAKSLKNKIWNBFYGGFSLGYSNTQALPPSGEVN
DAXKMKKHLKLAKSSADRFNLGRTMIGLDTNNYKWNKNGWRAISFSLVIN
NALLRANLQGENPSHYGITAFNPLNLTQSLSEVAXMTTSDVLDVLSICVIFAMSFY
PASVYFLQEVSRKAKHLQFISGVKPVYIWLNSFVDMCNVVPATLIIIFICFOQ
KSYVSTNLPLALLLLYGSITFLMPASFVFKIPSTAYVVLVSNI.FIGINGSA
TFVLELFTDKLNNINDILKSVLFIPLPCLAGRLIDVKNQAMADALRFGENRFS
P.LSWDLVGRNLFAMAVEGVVFLITVLIQYRFRIRPVPYNAKLSINDEDEVRERQ
RLDGGGQNDILEIKELTKYRKKKAPVDRCYGIPGCEGFLIGVAGCAGKSTFKM
LIGDITVRCGDAFLXANSILSNLHEVHQNGICPQFDATELUTGREHVEFFALLRGV
PEKEKGVEGWAIRKIGLVYKGYKAGNYSGGNKKRLSTAMALIGGPPVVFDEPITG
MDPKARFLWNCALSVYKGRSVLTSMSHECEALCTRMAIIVNGRFLGSLGVQHLK
NRFSGDGTIVRIAGSNPDLKPVQDFGLAFTPGSVYKKEKRNMLQVLPSSLSLARI
FSILSQSKRLHIEDYSVSQTTLDOVFVNFQAKDQSDDDHLKDLSLHKNTVVDVAULT
SFLQDEKVKESYV"
29916. .30150
/rpt_family="Alu"
31050. .31314
/rpt_family="Alu"
31692. .31973
/rpt_family="Alu"
36167. .36440
/rpt_family="Alu"
37617. .37892
/rpt_family="Alu"
37999. .38166
/rpt_family="Alu"
38172. .38269
/rpt_family="Alu"
38995. .39271
/rpt_family="Alu"
41547. .41565
/gene="ABCA1"
/ncore="A009X28"
41593. .41835
/rpt_family="Alu"
41657. .41679
/gene="ABCA1"
/ncore="A009X28"
42320. .42584
/rpt_family="Alu"
44349. .44626
/rpt_family="Alu"
48473. .48669
/rpt_family="Alu"
50547. .50614
/rpt_family="Alu"
54285. .54418

```


ROVMAEYVNTKTOELAVFHLDECMWELSPKWTWPMENSOEMDLVRMLLDSRDNDHEWE
COLDC-DRTAQDIAVFLAK#EDVQSSNGSVYTWREAFNETNOAIRCT-SRMECVNLN
KLEPATIEWLWLNKSNELLDERKFWAGVFGVIGPGSIELPHVYKJRMEDINDVERT
NKIDGYWDPGRADPFEDMRVWGCFAYLQDVEQALRVLTGTEKKTGYVMQGPY
PCYVDIFLRNRSMPFLFWLATYSAVVIKGIYEKEARLRETRIMRLDNLISILW
PSWFISSILPILVYAGLLVILKGNLLPSPDPSVYFVLSYFAVTLQCFLLISTILF
SRANLAACGGIIEFTLPLPYCLVWQDYGFTLKFASLLSPVAFGFCBCFALFE
EOGIGVQWQDLPFESVEEDGNLTTSVSMMLPNTFLYGVMTWYIEAFPGQYIPRPW
YPCCKSYWFESEDEKSEPGSNOKRISEICMEEPETHLKLGVSONLVKVRDGMKV
AVDGLALNFBQIHSIFLHNCAGAKTITMSILGLFPPTSGTAYILGXDIRSEMSIIR
QNLGVQPNVULFMTVEEHWFYARLKLGSKKHVKAEOMEQALDVGLPSSKLASKT
SOLSCGMQRKSVAILAFVGGKSVWILDEPTAGVDYPSRGIWELLKYRQGTIILST
HMDADYLDGRIALISHKLCVGSJFLKNQLGTYLTLVKKDVSSLSSCRNS
SVYLLKKDSVQSSDAGLGSDBESDTLTDVSAISNLRKHVSEARLVBDIGHEL
TVLPEAKGAFVFLHEIDRLSLGSSVGETTLEBIFLKVABESVDIAETS
DGTLPARNRFRAGKQCLRPETEDDAADPNDSIDPESRETDLLSGMDGKSYOVK
GWLKQOQFVALLKRLILARRSRKQFTQAVLPAFVFCIALVFLIIVPPFKYSLE
LQPMWNYBOYTVSDAPEDTGTLELLNALTKDPGFGTRCMEGNDIPDTPCQAGBEEW
TTAPVPQITIMDLFQNGWNTMQN#SPACOSSDKIKKMLPVPFGAGLPPQKQNTA
DILQDITNSIDYLVKTYVQIIAKSLXNKIWNFRYGGFSLGVSNQALPSPQFN
DAIKOMKHLKAKDSADRPLNSLRPMIGLDRNNKVMWNNKGWHAISFLFLAVIN
NALRANLOKGNPSHYGITAFNHPNLTKOOLSEVALMTTSDVILVSCVIFAKSFV
PASVYVFLQERVSRAKELQFISGVKPYIYNESNFWDMCNVVPATILVITIPQOQ
KSYVSTNLFVJALLLLYGSWITPMPASVFKIPSTAYVYVLTSVNLFINGNSVA
TFVLELTDNKLNNINDILKSVLFIHPHFCGLGRGLDMVKQAMADALERGENRVS
PLSWDLGNLFAMAVEGVVFFLITVLYQYREFIRPVPYNAKLSPLNDEDEVRREKQ
RLDGGQNDILEIKELTKIYRKRKPAYDRICVGPCEGFCGLLVNAGAKSYFKM
LTGDTWY3GDAFLAKNSILSNHEVHONWGYCPOPDATTELLTGREHVEFTAILRGV
PEKEVGKGEWAIKRLGLVKYGEKYAGNYSGGNKKKLSTAMALIGPPVYVDEPTIG
MDPKARFLWNCALSVMKGRSVLITSHMEECEALCTRNAMVNGFRCGLGSVOELK
NREGDGYTIVRIAGSNPLKXPQDFGLAFPGSVLKEXHRMLQYQLPSSLSLARI
FSLQSKRKLHIEDYVSQITLDQVYVNAFKAQDQDDHKLKDLSLHKNQTVYDVAVLI
SFIQDEKVKESV"
72893..72986
/gene="ABCA1"
/number=3
77520..77661
/gene="ABCA1"
/number=4
78931..79049
/gene="ABCA1"
/number=5
101348..101169
/gene="ABCA1"
/number=6
104152..104328
/gene="ABCA1"
/number=7
117287..117379
/gene="ABCA1"
/number=8
122329..122569
/gene="ABCA1"
/number=9
125285..125424
/gene="ABCA1"
/number=10
125754..125870
/gene="ABCA1"
/number=11
130579..130276
/gene="ABCA1"
/number=12
131023..131228
/gene="ABCA1"
/number=13
131749..131925
/gene="ABCA1"
/number=14
133708..133930
/gene="ABCA1"
/number=15
135677..135899

/gene="ABCA1"
/number=16
136959..137163
/gene="ABCA1"
/number=17
138268..138381
/gene="ABCA1"
/number=18
140179..140350
/gene="ABCA1"
/number=19
14340..141471
/gene="ABCA1"
/number=20
142777..142919
/gene="ABCA1"
/number=21
143123..143260
/gene="ABCA1"
/number=22
143963..144183
/gene="ABCA1"
/number=23
145441..145513
/gene="ABCA1"
/number=24
146657..146859
/gene="ABCA1"
/number=25
148527..148575
/gene="ABCA1"
/number=26
148771..148884
/gene="ABCA1"
/number=27
150280..150428
/gene="ABCA1"
/number=28
152078..152202
/gene="ABCA1"
/number=29
153438..153536
/gene="ABCA1"
/number=30
156568..156757
/gene="ABCA1"

Query Match 95.4%; Score 441.6; DB 9; Length 201144;
Best Local Similarity 98.1%; Pred. No. 1.8e-89;
Matches 459; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 1 CCTTGGCTGCGGGAAAGTGGACTAGAGAGTCTGCGGCGAGCCCGAGCCGAGCCGCGTTC 60
DB 33554 CC-TGGGTGCGGGAAAGTGGACTAGAGAGTCTGCGGCGAGCCCGAGCCGAGCCGCGTTC 33613

QY 61 CCGCGCTCTTTAGCCCGCGGCGCCCGGGGGGGAAGGACCGAGACCGGGGACCCCTAA 120
DB 33614 CCGCGCTCTTTAGCTGGCGGCGCCCGGGGGAAGGACCGAGACCGCGGACCCCTAA 33673

QY 121 GACACCTGCTGTACCTCCAC-CC-CCACACCCACCCACCTCCCGCCCACTCCCTAGA 175
DB 33674 GACACCTGCTGTACCTCCACCCACCCACCCACCTCCCGCCCACTCCCTAGA 33733

QY 176 TGTGTCGTGGCGGCTGAAGCTGCGCCGTTTAAAGGGCGGCGCCCGGCTCCACGTGCTTT 235
DB 33734 TGTGTCGTGGCGGCTGAAGCTGCGCCGTTTAAAGGGCGGCGCCCGGCTCCACGTGCTTT 33793

QY 236 CTGCTGAGTACTGAACCTACATAAACAGAGCCCGGGAACCGGGGCGGGAGGAGGAGAGC 295
DB 33794 CTGCTGAGTACTGAACCTACATAAACAGAGCCCGGGAAGGAGGCGGGGAGGAGGAGAGC 33853

QY 296 ACAGGCTTGCCGATAGTACCTCTGGGCTCGGTGCGCGCGGANTCTATATAAGAACTA 355
|||||

```

Db 33854 ACAGGCTTTGACCGATAGTAAACCTCTCGCTCGTGCAGCCGAATCTATAAAAGGAACTA 33913
QY 356 GTCCGGCCAAAACCCCGTAATTCGCGAGCGAGAGTAGTGGCGCGGACCCGACGAGCC 415
|||||
Db 33914 GTCTCGCAAAAACCCCGTAATTCGCGAGCGAGAGTAGTGGCGCGGACCCGACGAGCC 33973
QY 416 GAGCGCACCCCTTCTCTCCCGGGCTGGCGCAGGCGAGGCGGGAGCTC 463
|||||
Db 33974 GAGCGCACCCCTTCTCTCCCGGGCTGGCGCAGGCGAGGCGGGAGGCTC 34021
|||||

RESULT 12
AX351031 AX351031 2893 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 3 from Patent WO0183746.
DEFINITION
ACCESSION AX351031
VERSION AX351031.1 GI:18616387
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Benefie, P.,
Stewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
Regulatory nucleic acid sequences of the abcl gene
Patent: WO 0183746-A 3 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
Source 1..2893
Location/Qualifiers
1..2893
Query Match 76.88; Score 355.4; DB 6; Length 2893;
Sest Local Similarity 98.4%; Pred. No. 6.6e-70;
Matches 371; Conservative 0; Mismatches 1; Indels 5; Gaps 1:
QY 1 CTTTGGCTGCGCGGAACCTGGACTAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTC 60
|||||
Db 2517 CTTTGGCTGCGCGGACGTGGACTAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTC 2576
|||||
QY 61 CCGGCGCTTTAGCGCGCGCGCGCGCGGCGGGAAGGGACGCGAGCCGCGGACCCCTAA 120
|||||
Db 2577 CCGGCGCTTTAGCGCGCGCGCGCGCGGCGGGAAGGGACGCGAGCCGCGGACCCCTAA 2636
|||||
QY 121 GACACTGCTCTACCTCCACCCCGAC-----CCACCCACCTCCCGCCCACTCCCTAGA 175
|||||
Db 2637 GACACTGCTCTACCTCCACCCCGACCCCGACCCACCTCCCGCCCACTCCCTAGA 2695
|||||
QY 176 TGTGTCGTGGCGGCTGACCTGCGCCCGCTTTAAGGGCGCGGCGCGGCTCCACGTCGTTT 235
|||||
Db 2697 TGTGTCGTGGCGGCTGACCTGCGCCCGCTTTAAGGGCGCGGCGCGGCTCCACGTCGTTT 2756
|||||
QY 236 TGTGTCGTGACTGACTACATACATACAGCGCCGCGGAGCGGGCGGGGAGGAGGAGGAGC 295
|||||
Db 2757 TGTGTCGTGACTGACTACATACATACAGCGCCGCGGAGCGGGCGGGGAGGAGGAGGAGC 2816
|||||
QY 296 ACAGGCTTTGACCGATAGTAACTCTGCGCTGCGGTGCGAGCCGCAACTATATAAGGAACTA 355
|||||
Db 2817 ACAGGCTTTGACCGATAGTAACTCTGCGCTGCGGTGCGAGCCGCAACTATATAAGGAACTA 2876
|||||
QY 356 GTCCCGGCAAAAACCCC 372
|||||
Db 2877 GTCCCGGCAAAAACCCC 2893
|||||

RESULT 13
AC021246 AC021246/c 69570 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens clone R211-1N10, LOW-PASS SEQUENCE SAMPLING.
DEFINITION
ACCESSION AC021246

```

VERSION KEYWORDS SOURCE ORGANISM

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS

```

AC021246.2 GI:9119882
HTG: HTGS_PEA50.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69570)
Birren, B., Linton, L., Nusbaum, C. and Lander, B.
Homo sapiens chromosome, clone RP11-1N10
Unpublished
2 (bases 1 to 69570)
Birren, B., Linton, L., Nusbaum, C., Lander, B., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavskiy, L., Bouckgalter, B., Brown, A., Burkett, G., Castle, A.,
Choepl, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeArallano, K., Dewar, K., Domino, M., Doyle, M., Fenesfor, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karates, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, J.M., Peterson, K.,
Piere, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severi, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.C.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://p.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: L_N_10
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 871: contig of 871 bp in length
372 971: gap of 100 bp
972 1834: contig of 863 bp in length
1835 1934: gap of 100 bp
1935 2804: contig of 870 bp in length
2805 2904: gap of 100 bp
2905 3745: contig of 841 bp in length
3746 3845: gap of 100 bp
3846 4696: contig of 851 bp in length
4697 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 6540: contig of 800 bp in length
6541 6640: gap of 100 bp
6641 7509: contig of 865 bp in length
7510 7609: gap of 100 bp
7610 8479: contig of 870 bp in length
8480 8579: gap of 100 bp

```

TITLE JOURNAL COMMENT


```

Db 67402 AGCGAGATGAGTGGCGGCGGACCGACAGCAGCGGACCTCTCTCCGGCTGCG 67343
|||||
QY 442 GGCGAGGCGAGGCGCGGAGGTC 463
|||||
Db 67342 GACAGGACAGGCGGCGGAGGTC 67321
|||||

RESULT 14
AC021345/c
LOCUS AC021345
DEFINITION Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90698)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavsky, I., Boukhvalter, B., Brown, A., Burkett, G., Castle, A.,
Choquel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeArrellano, K., Dewar, K., Domino, M., Doyle, M., Fenescor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., LeCock, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneses, L., Morrow, J., Naylot, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severi, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Vlei, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4483
Center clone name: 24_J_9
-----
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length

```

```

2825 2924: gap of 100 bp
2925 3802: contig of 878 bp in length
3803 3902: gap of 100 bp
3903 4816: contig of 914 bp in length
4817 4916: gap of 100 bp
4917 5759: contig of 843 bp in length
5760 5859: gap of 100 bp
5860 6764: contig of 905 bp in length
6765 6864: gap of 100 bp
6865 7747: contig of 883 bp in length
7748 7847: gap of 100 bp
7848 8755: contig of 908 bp in length
8756 8855: gap of 100 bp
8856 9753: contig of 898 bp in length
9754 9853: gap of 100 bp
9854 10757: contig of 904 bp in length
10758 10857: gap of 100 bp
10858 11732: contig of 875 bp in length
11733 11832: gap of 100 bp
11833 12739: contig of 907 bp in length
12740 12839: gap of 100 bp
12840 13710: contig of 871 bp in length
13711 13810: gap of 100 bp
13811 14684: contig of 874 bp in length
14685 14784: gap of 100 bp
14785 15662: contig of 878 bp in length
15663 15762: gap of 100 bp
15763 16677: contig of 915 bp in length
16678 16777: gap of 100 bp
16778 17678: contig of 901 bp in length
17679 17778: gap of 100 bp
17779 18679: contig of 901 bp in length
18680 18779: gap of 100 bp
18780 19632: contig of 853 bp in length
19633 19732: gap of 100 bp
19733 20634: contig of 902 bp in length
20635 20734: gap of 100 bp
20735 21620: contig of 886 bp in length
21621 21720: gap of 100 bp
21721 22578: contig of 859 bp in length
22580 22679: gap of 100 bp
22680 23568: contig of 889 bp in length
23569 23668: gap of 100 bp
23669 24554: contig of 886 bp in length
24555 24654: gap of 100 bp
24655 25521: contig of 867 bp in length
25522 25621: gap of 100 bp
25622 26487: contig of 866 bp in length
26488 26587: gap of 100 bp
26588 27464: contig of 877 bp in length
27465 27564: gap of 100 bp
27565 28466: contig of 902 bp in length
28467 28566: gap of 100 bp
28567 29454: contig of 898 bp in length
29455 29564: gap of 100 bp
29565 30447: contig of 883 bp in length
30448 30547: gap of 100 bp
30548 31453: contig of 906 bp in length
31454 31553: gap of 100 bp
31554 32452: contig of 899 bp in length
32453 32552: gap of 100 bp
32553 33447: contig of 895 bp in length
33448 33547: gap of 100 bp
33548 34435: contig of 888 bp in length
34436 34535: gap of 100 bp
34536 35433: contig of 898 bp in length
35434 35533: gap of 100 bp
35534 36440: contig of 907 bp in length
36441 36540: gap of 100 bp
36541 37422: contig of 882 bp in length
37423 37522: gap of 100 bp
37523 38402: contig of 880 bp in length
38403 38502: gap of 100 bp

```

```
* 38503 39380: contig of 878 bp in length
* 3938: 35480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp
```

Query Match 49.6% Score 229.8; DB 2; Length 90698;
Best Local Similarity 56.0%; Pred. No. 7.8e-42;
Matches 279; Conservative 0; Mismatches 141; Indels 3; Gaps 3;

QY 41 AGCCCGGAGCCACCGCTCCCGCGCGCTCTAGCGCCGCGCGCCGCGCGGCGGCGGGAAGGG 100

|||||

Db 87653 AGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 87604

```
QY 101 ACGCAGACGGCGGACCTAAGACACCTGCTGTACCTCCACCCCAACCCACCCACCTCC 160
Db 87803 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 57744
QY 161 CCCAACCTCCCTAGATGCTGCTGGCGGCTGAAAGTCCCGCTTTAAGGGGCGGCGCC 220
Db 87743 CCACCTCCCGCACTATAGTGTGCTGCGCTGAAGCTGCGCGG-TTAAAGGGGCGGCGCC 87685
QY 221 GGTCCACCGTCTTCTGCTGAGTGAACCTACATAAAGAGCGCGGGAACGGGCG 280
Db 87684 GGTCCACCGTCTTCTGCTGAGTGAACCTACATAAAGAGAGAGAGCGGCGGCG 87626
QY 281 GGGAGAGGGGAGAGACAGAGCTTTGACCGATAGTAACTCTGCTGCTGCTGAGCGCAAT 340
Db 87625 GGGAGAGGGGAGAGACAGAGCTTTGACCGATAGTAACTCTGCTGCTGCTGAGCGCAAT 87566
QY 341 CTATAAAGGAAGTCTGCGGCAAAACCCGTAATTCGGACGAGAGTGAAGTGGGCG 400
Db 87565 CTATAAAGGAAGTCTGCGGCAAAACCCGTAATTCGGACGAGAGTGAAGTGGGCG 87507
QY 401 GGGACCGCGAGAGCGGACCGCTTCTCTCCCGGCTGGCGAGGCGGCGGCGGAG 460
Db 87506 GGGACCGCGAGAGCGGACCGCTTCTCTCCCGGCTGGCGAGGCGGCGGCGGAG 87447
QY 461 CTC 463
Db 87446 CTC 87444

RESULT 15
AF258627 Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
LOCUS partial cds.
DEFINITION AF258627.1 GI:7759707
ACCESSION AF258627
VERSION AF258627.1
KEYWORDS .
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.C. and Kane,J.P.
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.C. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
FEATURES
source
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABCI"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AA569513.1"
/db_xref="GI:7759708"
/translation="MACWQQLRLLLKLNLTFRPOTCOLLLEVAWPLFLILISYRL
SYPPYSCHECHFPKAMPAGTLPVVOGITCNANNPCRYPTPGAPGVGVGNFKS"
```

```
BASE COUNT      152 a      198 c      190 g      156 t      1 others
ORIGIN

Query Match      37.4%; Score 173; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. Nc. 7.7e-29;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 AGAGCACAGGCTTTGACCGATAGTACCTCTGCGCTGCGGTGCAGCCGAATCTATAAAGG 350
Db 1 AGAGCACAGGCTTTGACCGATAGTACCTCTGCGCTGCGGTGCAGCCGAATCTATAAAGG 60
QY 351 AACTAGTCCCGGCAAAACCCCGTAATTGCGAGCGAGAGTGAAGTGGCGCGGACCCGCA 410
Db 61 AACTAGTCCCGGCAAAACCCCGTAATTGCGAGCGAGAGTGAAGTGGCGCGGACCCGCA 120
QY 411 GAGCCGAGCGGACCTTCTCTCCGGGCTCCGGGAGGGGAGGAGGTC 463
Db 121 GAGCCGAGCGGACCTTCTCTCCGGGCTCCGGGAGGGGAGGAGGTC 173
```

Search completed: February 4, 2003, 01:26:54
Job time : 1675.59 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:39:45 ; Search time 876.075 Seconds
(without alignments)
8559.21C Million cell updates/sec

Title: US-09-596-141c-3_COPY_1181_1643
Perfect score: 463
Sequence: 1 ccttgctgcgggaacgtg.....caggcgaggcgaggagctc 463

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 805774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfn:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89.4	19.3	736	9	AU135588
2	89	19.2	547	9	AL698654
3	89	19.2	763	9	AU121731
4	75	16.2	292	14	Z44377
5	75	15.2	1004	11	BC034824
6	60.2	13.0	998	12	BG678861

7	55.2	11.9	1201	17	CNS015W3
8	53.8	11.6	298	17	CNS04M2X
9	50.6	10.9	1101	17	CNS017SY
10	48.2	10.4	925	17	CNS0091P
11	47.6	10.3	870	17	CNS006EZ
12	47.6	10.3	1009	17	CNS010EW
13	47.6	10.3	1101	17	CNS01720
14	47.4	10.2	1101	17	CNS017V2
15	47	10.2	843	17	CNS00C51
16	46.2	10.0	681	17	CNS02FE9
17	46.2	10.0	900	17	AG081217
18	46.2	10.0	925	17	CNS0091P
19	46.2	10.0	1136	14	BQ943816
20	46	9.9	758	17	AG073529
21	46	9.9	1004	14	BQ544204
22	45.8	9.9	523	9	AL514015
23	45.8	9.9	997	14	BQ553024
24	45.6	9.8	456	9	AL554951
25	45.6	9.8	805	17	AG157500
26	45.6	9.8	1213	13	BM560246
27	45.4	9.8	766	17	AG041031
28	45.2	9.8	853	14	BQ219067
29	45.2	9.8	875	17	AG043475
30	45	9.7	883	12	BG086652
31	45	9.7	884	17	CNS018NP
32	45	9.7	1139	14	BM914614
33	45	9.7	1169	13	BM549879
34	44.8	9.7	875	17	AG043475
35	44.8	9.7	1150	14	BQ892482
36	44.8	9.7	1153	17	AG065289
37	44.2	9.5	477	9	AL513813
38	44.2	9.5	456	17	CNS00K3H
39	44.2	9.5	927	17	CNS04W3S
40	44.2	9.5	932	17	CNS0072C
41	44.2	9.5	942	17	AG042878
42	44.2	9.5	1059	14	BQ680479
43	44	9.5	403	9	AL513741
44	44	9.5	694	17	AG060229
45	44	9.5	810	12	BF262468

ALIGNMENTS

RESULT 1
AU135588
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AU135588 PLACE1 Homo sapiens cDNA clone PLAC1002437 5', mRNA
sequence.
AU135588
EST.
human.
Homo sapiens
Dukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 736)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0512, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

```

FEATURES
  source      Location/Qualifiers
  1..736
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="PLACE1002437"
    /clone_lib="PLACE1"
    /tissue_type="placenta"
    /note="vector: pME18SPL3"
  BASE COUNT      163 a   199 c   199 g   170 t      5 others
  ORIGIN

Query Match      19.3%; Score 89.4; DB 9; Length 736;
Best Local Similarity 98.9%; Pred. No. 1.2e-11;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 373 GTAATTCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 432
Db 1 GTAGTTCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 60
|||||
QY 433 CCGGCTCGCGAGCGAGCGCGGCGGAGCTC 463
Db 61 CCGGCTCGCGAGCGAGCGCGGCGGAGCTC 91
|||||

RESULT 2
LOCUS      AL698654      547 bp      mRNA      linear      EST 21-MAR-2002
DEFINITION Dkf2p686N12109_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION Dkf2p686N12109 5', mRNA sequence.
VERSION AL698654
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1995)
COMMENT Contact: Duesterhoeft A
MIPS Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686N12109) is available at the R2PD in Berlin.
Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzd.ge.
FEATURES
  source      Location/Qualifiers
  1..547
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="DKFZp686N12109"
    /clone_lib="686 (synonym: hlcc3)"
    /tissue_type="human skeletal muscle"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
    cDNA-collection"
  BASE COUNT      112 a   154 c   157 g   124 t
  ORIGIN

Query Match      19.2%; Score 89; DB 9; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 AATTCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 434
|||||

```

```

Do 2 AATTGCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 61
QY 435 GGGCTCGGCGAGCGCGGCGGAGCTC 463
|||||
Db 62 GGGCTCGGCGAGCGCGGCGGAGCTC 90
|||||

RESULT 3
LOCUS      AU121731      763 bp      mRNA      linear      EST 01-AUG-2002
DEFINITION AU121731 MAMMAL Homo sapiens cDNA clone MAMMAL000851 5', mRNA
sequence.
ACCESSION AU121731
VERSION AU121731.1 GI:10936966
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Kakamura,Y., Nagai,T., Sugano,S. and
Isogai,I.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana-Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
  source      Location/Qualifiers
  1..763
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="MAMMAL000851"
    /clone_lib="MAMMAL"
    /tissue_type="mammary gland"
    /note="vector: pME18SFL3"
  BASE COUNT      137 a   205 c   260 g   158 t      3 others
  ORIGIN

Query Match      19.2%; Score 89; DB 9; Length 763;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 AATTGCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 434
|||||
Db 1 AATTGCGAGCGAGTGAGTGGGGCGGAGCCGCGAGCGCGAGCCCTCTCTCC 60
|||||

QY 435 GGGCTCGGCGAGCGCGGCGGAGCTC 463
|||||
Db 61 GGGCTCGGCGAGCGCGGAGCTC 89
|||||

RESULT 4
LOCUS      244377      292 bp      mRNA      linear      EST 14-NOV-1994
DEFINITION HSC12B081 normalized infant brain cDNA Homo sapiens cDNA clone
c-12b08, mRNA sequence.
ACCESSION 244377
VERSION 244377.1 GI:573506
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 292)

```

AUTHORS

Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Piestu, G., Pouliot, I., Sebastiani-Kabakchis, C., and Tessier, A.

TITLE

IMAGE: molecular integration of the analysis of the human genome and its expression

JOURNAL

C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

MEDLINE

95277534

COMMENT

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr

Single read

Genexpress_library_id: C; Genexpress_sequence_id: y-c-1zb08
Seq primer: (-21)M13-universal.

FEATURES

source

1. .292
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-1zb08"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafamid BA; Site:1: HindIII;
Site:2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafamid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 96 g 56 t 3 others
ORIGIN
Query Match 16.2%; Score 75; DB 14; Length 292;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 389 GTAGTGGGCGGGACCCGAGCGAGCGAGCCGACCGCTTCTCTCCGGGCTGGCGCAGGG 448
|||||
Db 1 GTAGTGGGCGGGACCCGAGCGAGCGAGCCGACCGCTTCTCTCCGGGCTGGCGCAGGG 60
QY 449 CAGGGCGGGGAGCTC 463
|||||
Db 61 CAGGGCGGGGAGCTC 75

RESULT 5

BC034824

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC034824 1004 bp mRNA linear HTC 26-JUL-2002
Homo sapiens, clone IMAGE:4749735, mRNA.
BC034824
BC034824.1 GI:21961568
HTC.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgebps-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amc@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate. 52 Row: d Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5915657
This clone has the following problem: incomplete processing.

FEATURES

source

1. .1004
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749735"
/tissue_type="Skin, squamous cell carcinoma"
/clone_lib="NCI_CGAP_Skn4"
/lab_host="DH103"
/note="Vector: pCMV-SPORT6.cdb"

BASE COUNT 306 a 234 c 230 g 234 t
ORIGIN
Query Match 16.2%; Score 75; DB 11; Length 1004;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 389 GTAGTGGGCGGGACCCGAGCGAGCGAGCCGACCGCTTCTCTCCGGGCTGGCGCAGGG 448
|||||
Db 1 GTAGTGGGCGGGACCCGAGCGAGCGAGCCGACCGCTTCTCTCCGGGCTGGCGCAGGG 60
QY 449 CAGGGCGGGGAGCTC 463
|||||
Db 61 CAGGGCGGGGAGCTC 75

RESULT 6

BG578861

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG578861 998 bp mRNA linear EST 01-MAY-2001
602624750F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5',
mRNA sequence.
BG578861
BG578861.1 GI:13910258
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgebps-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM0503 row: g column: 16
High quality sequence stop: 860.

FEATURES

source

1. .998
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/clone="IMAGE:4749735"
/clone_lib="NCL_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DHI0B (T1 phage-resistant)"
/note="Organ: skin; vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCL_CGAP Library."

BASE COUNT      285 a   233 c   244 g   236 t
ORIGIN

Query Match      13.0%; Score 60.2; DB 12; Length 998;
Best Local Similarity 89.0%; Pred. No. 0.00015;
Matches 65; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 391 GAGTGGCGCGGACCGCAGAGCGAGCGACCGCTTCTCCCGGGCTCGCGGAGGGCA 450
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1  GTAGTGGCGCGGACCGCAGAGCGAGCGACCGCTTCTCCCGGGCTCGCGGAGGGCA 60

QY 451 GGGCGGGGAGCTC 463
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 GGGCGGGGAGCTC 73

RESULT 7
CNS015W3      CNS015W3      1201 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence S96 end of BAC
DEFINITION      BAC15123 of DrosBAC library from drosophila melanogaster (fruit
      fly), genomic survey sequence.
ACCESSION      AL105981
VERSION      AL105981.1 GI:5619397
KEYWORDS
SOURCE      Drosophila melanogaster.
ORGANISM      Drosophila melanogaster.
      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
      Ephydroidea; Drosophiliidae; Drosophila.
REFERENCE      1 (bases 1 to 1201)
      Genoscope.
      Direct Submission
      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
      BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
      - Web : www.genoscope.cns.fr)
      Determination of this BAC-end sequence was carried out as part of a
      collaboration with the European Drosophila Genome Project (EDGP) -
      http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
      library (Dros BAC) was made by Alain Billaut at CEPE (Centre
      d'Etude du Polymorphisme Humain) with funding provided by a MRC
      project grant. The DNA was prepared from embryos by Alain Bucheton
      and Genevieve Payan. It has been constructed in the vector
      pBelOBAC11.

FEATURES
      source      Location/Qualifiers
      1..1201
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="BACN15123"
      /clone_lib="DrosBAC"
      /plasmid="pBelOBAC11"
      /notes="end : S96"

BASE COUNT      255 a   276 c   235 t   157 others
ORIGIN

Query Match      11.9%; Score 55.2; DB 17; Length 1201;
Best Local Similarity 41.0%; Pred. No. 0.0024;
Matches 91; Conservative 34; Mismatches 97; Indels 0; Gaps 0;

QY 5  GGTGCGCGGACGTGCGACTAGAGAGTGTGGCGSCAGCCCGAGCCGACGCTTCCCGC 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 927 GCGCGCTCCGCGSCCARASMGGGSSGGGCGCCMGCGCGSCNGCGCGGSCCGSC 986

QY 55 GGTCTTAGGCGGCGCGCGCGGAGGAGGCGGAGCGGAGCGGAGCCCTTAAGACA 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 987 CNGCSSSSSGGGGGGGMCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1046

QY 125 CTGCTGTACCTTCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1047 CCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCSG 1106
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 GCGGCTGAACGTCGCGCGCTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 226
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1107 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1148

RESULT 8
CNS04M2X      CNS04M2X      298 bp      DNA      linear      GSS 21-MAY-2000
LOCUS      Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION      119N17 of library G from Tetraodon nigroviridis, genomic survey
      sequence.
ACCESSION      AL296898
VERSION      AL296898.1 GI:8035478
KEYWORDS      GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis.
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
      Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
      Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 298)
      Roest-Crolius H., Jaillon O., Dasilva C., Bouneau L., Fisher C.,
      Bernot A., Fizames C., Wincker P., Brottier P., Quetier F.,
      Saurin W. and Weissenbach J.
      Human gene number estimate provided by genome wide analysis using
      Tetraodon nigroviridis DNA sequence
      Unpublished
REFERENCE      2 (bases 1 to 298)
      Roest-Crolius H., Jaillon O., Dasilva C., Fizames C., Fisher C.,
      Bouneau L., Billault A., Quetier F., Saurin W., Bernot A. and
      Weissenbach J.
      Characterization and repeat analysis of the compact genome of the
      freshwater pufferfish Tetraodon nigroviridis
      Unpublished
REFERENCE      3 (bases 1 to 298)
      Genoscope.
      Direct Submission
      Submitted (12-APR-2000)
      This sequence is a single read and was generated as part of a large
      scale clone-end sequencing project of the Tetraodon nigroviridis
      genome. For more information, please take a look at
      http://www.genoscope.cns.fr/tetraodon.
      Location/Qualifiers
      1..298
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone="119N:7"
      /clone_lib="G"

BASE COUNT      18 a   106 c   80 g   15 t   79 others
ORIGIN

Query Match      11.6%; Score 53.8; DB 17; Length 298;
Best Local Similarity 40.2%; Pred. No. 0.0047;
Matches 80; Conservative 33; Mismatches 86; Indels 0; Gaps 0;

QY 1  CCTTGGCTGCGGGAACGTGACTAGAGTGTGGGGCGAGCCCGGACCGCGCTTC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 83 CNKSCCVDVGGCCVCGSGSGSGSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSG 142

QY 61 CCGCGCGCTTCTAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 143 GSGSCSCSGSGSGSGSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202

QY 121 GACACCTGTCTACCTTCCACCCACCCACCCACCCACCCACCCACCCACCTCCCTAGATGTGT 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 203 CNCNCNCCTCNCNCNCNNNNCCCCCCCCCCCCCCCCCCCCCCCCCCCCGGGGGG 262

```



```
Search completed: February 4, 2003, 03:26:33
Job time : 888.789 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:56:40 ; Search time 21.7309 Seconds
(without alignments)
6534.066 Million cell updates/sec

Title: US-09-596-141c-3_COPY_1181_1643
Perfect score: 463
Sequence: 1 cctggctgcgggaacgtg.....cagggcagggcgggagctc 463

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2.6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.6	8.3	4481	US-08-041-886-18	Sequence 18, Appl
2	38.4	8.3	595	US-08-483-533-4	Sequence 4, Appl
3	38.4	8.3	595	US-08-283-471A-4	Sequence 4, Appl
4	38.4	8.3	1327	US-08-483-533-36	Sequence 36, Appl
5	38.4	8.3	1327	US-09-283-471A-36	Sequence 36, Appl
6	38.4	8.3	1335	PCT-US91-06532-1	Sequence 1, Appl
7	37.6	8.1	936	US-08-018-977C-4	Sequence 4, Appl
8	36.2	7.8	43280	US-08-804-227C-1	Sequence 1, Appl
9	35.8	7.7	4085	US-08-072-574-7	Sequence 7, Appl
10	35.8	7.7	4181	US-08-072-574-9	Sequence 9, Appl
11	35.8	7.7	11907	US-08-061-376-4	Sequence 4, Appl
12	35.6	7.7	4411529	US-09-103-840A-1	Sequence 1, Appl
13	35.6	7.7	4411529	US-09-103-840A-1	Sequence 1, Appl
14	35.2	7.6	35100	US-08-306-694B-19	Sequence 19, Appl
15	35.2	7.6	35100	PCT-US93-05251-19	Sequence 19, Appl
16	35	7.6	7218	US-08-232-463-14	Sequence 14, Appl
17	34.6	7.5	1280	US-08-483-533-38	Sequence 38, Appl
18	34.6	7.5	1280	US-09-283-471A-38	Sequence 38, Appl
19	34.6	7.5	17410	US-07-841-646-3	Sequence 3, Appl
20	34.6	7.5	17410	US-08-147-023-3	Sequence 3, Appl
21	34.6	7.5	17410	US-08-447-570-3	Sequence 3, Appl
22	34.6	7.5	17410	US-08-449-700-3	Sequence 3, Appl
23	34.6	7.5	17410	US-08-449-699A-3	Sequence 3, Appl
24	34.4	7.4	954	US-08-170-294-1	Sequence 1, Appl
25	34.4	7.4	954	US-08-664-855-1	Sequence 1, Appl
26	34.4	7.4	954	US-09-049-289-1	Sequence 1, Appl
27	34.4	7.4	2888	US-08-765-907A-1	Sequence 1, Appl

C 28	34.2	7.4	1490	4	US-08-964-127-3	Sequence 3, Appl
C 29	34.2	7.4	1490	4	US-09-496-692-3	Sequence 3, Appl
C 30	34.2	7.4	1610	4	US-09-276-531-56	Sequence 56, Appl
C 31	34.2	7.4	2460	4	US-08-964-127-1	Sequence 1, Appl
C 32	34.2	7.4	2460	4	US-09-496-692-1	Sequence 1, Appl
C 33	34.2	7.4	4085	1	US-08-486-270-7	Sequence 7, Appl
C 34	34.2	7.4	4085	3	US-08-367-264-7	Sequence 7, Appl
C 35	34.2	7.4	4085	4	US-09-153-757-7	Sequence 7, Appl
C 36	34.2	7.4	4181	1	US-08-486-270-9	Sequence 9, Appl
C 37	34.2	7.4	4181	3	US-08-367-264-9	Sequence 9, Appl
C 38	34.2	7.4	4181	4	US-09-153-757-9	Sequence 9, Appl
C 39	34.2	7.4	4207	4	US-08-660-148-1	Sequence 1, Appl
C 40	34.2	7.4	4207	4	US-08-660-148-3	Sequence 3, Appl
C 41	34.2	7.4	4303	4	US-08-660-148-4	Sequence 4, Appl
C 42	34.2	7.4	4303	4	US-08-660-148-6	Sequence 6, Appl
C 43	34.2	7.4	72604	4	US-09-268-992-7	Sequence 7, Appl
C 44	34.2	7.4	72604	4	US-09-657-474-7	Sequence 7, Appl
C 45	33.8	7.3	2877	4	US-09-235-103-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-041-886-18
; Sequence 18, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabszadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163..4099
US-09-041-886-18

Query Match 8.3%; Score 38.6; DB 4; Length 4481;
Best Local Similarity 52.1%; Pred. No. 0.43; Indels 0; Gaps 0;
Matches 86; Conservative 0; Mismatches 79;
CY 7 CTGCGGGAACGTGGACTAGAGAGTCTGCGGCGCGAGCCGAGCCGAGCGCTTCCGCGCGC 66

Db	550	CCTGCGGCGCCTGCGCTCGAGCGCGCGCGAGGGCGCGCGAGCGCGCGCGAC	609
QY	61	CCGCGGCTCTTAGCGCGGCGGCGCGGGGAGGGAGCGACACGCGGAGCCCTAA	120
Db	610	CCCGGCGACCCCGCGACCCCGCGACCCCGGACCCCGCGACCCCGCGC	669
QY	121	GACACCTGCTGACCTCTCACGCCCGCAGCG--CACCGCACTCCGCGCAACTCCCTAGATG	177
Db	670	GACCCCGGAGACCCCGCGACCCCGCGCGGGTGTCTCTCGCCCACTCCGGGTGG	729
QY	178	TGTCGCTGGCGGCTGACCTGCGCCCGCTTTAAGGGGCGGCGCGGCTCCACCTGCTTCT	237
Db	730	CGACCTTGGTGTGTGGGCTCGGCGCCCGCGCTGTGGCGCGCGGCTCTGGCGCGCGA	789
QY	239	CTGAGTGACTGAAC	253
Db	790	GCGGGCGCACCGGGCT	805

```

1  RESULT 6
2  PCT-US91-06532-1
3  / Sequence 1, Application PC/TUS9106532
4  / GENERAL INFORMATION:
5  / APPLICANT: Roiztan, Bernard
6  / TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
7  / TITLE OF INVENTION: Vaccines and Methods
8  / NUMBER OF SEQUENCES: 8
9  / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
11 / ADDRESSES: Bicknell
12 / STREET: Two First National Plaza Suite 2100
13 / CITY: Chicago
14 / STATE: Illinois
15 / COUNTRY: USA
16 / ZIP: 60603
17 / COMPUTER READABLE FORM:
18 / MEDIUM TYPE: Floppy disk
19 / COMPUTER: IBM PC compatible
20 / OPERATING SYSTEM: PC-DOS/MS-DOS
21 / SOFTWARE: Patentic Release #1.0, Version #1.25
22 / CURRENT APPLICATION DATA:
23 / APPLICATION NUMBER: PCT/US91/06532
24 / FILING DATE: 19910910
25 / CLASSIFICATION: 424
26 / ATTORNEY/AGENT INFORMATION:
27 / NAME: Gruber, Lewis S.
28 / REGISTRATION NUMBER: 30,060
29 / REFERENCE/DOCKET NUMBER: 27373/8235
30 / TELECOMMUNICATION INFORMATION:
31 / TELEPHONE: 312/346-5750
32 / TELEFAX: 312/994-9740
33 / TELEX: 25-3856
34 / INFORMATION FOR SEQ ID NO: 1:
35 / SEQUENCE CHARACTERISTICS:
36 / LENGTH: 1335 base pairs
37 / TYPE: NUCLEIC ACID
38 / STRANDEDNESS: single
39 / TOPOLOGY: linear
40 / MOLECULE TYPE: DNA (genomic)
41 / PCT-US91-06532-1

```

```

Query Match      8.3%; Score 38.4; DS 5; Length 1335;
Best Local Similarity 49.6%; Pred. No. 0.39;
Matches 127; Conservative 0; Mismatches 126; Indels 3; Gaps
QY      1 CTTTGGCTGCCGGAAGCTGGACTAGAGATCTCTGGGCGCAGCCCGAGCCAGCCAGCCCTTC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      558 CTTGGCGCGCTGGAGCTTGGACGGCGGCGGCGGAGGGGGCGCCGGAGCCCCCGCGAC 617
QY      61 CCGGCGCTCTTAGCCCGCGGGGCGCCGGGCGGGGAGAGGAGCCAGACCGCGGAGCCCTAA 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      618 CCGCGGACCCCGCGACCCCGCGACCCCGCGAGATCCCGCGAGCCCGCGAGCCCGGACCCCGG 177

```

QY	121	GACACC	TGCTGTGGGGCGGCGTGGCCCGGCGGCCTTTAAAGSGGGGGGGCCCCGGGCTCCACGTGGCTTCT	177
Db	678	GACCCC	GGGAGCAGCCGCCGACGCCGGGGGGTGCGCTTCCTGCCGCCACGTCGCGGGTGCG	737
QY	178	TGCTGTGGGGCGGCGTGGCCCGGCGGCCTTTAAAGSGGGGGGGCCCCGGGCTCCACGTGGCTTCT	177	
Db	738	CACACTGTGTGGTGGGCCCGGCGCGCGCGCGCCCTGGGCGCGCGGCGCTCGTGGGCGCCCGCA	797	
QY	238	GTGTGAGTCACTGAAC	253	
Db	798	GCAGGCGCACCGGCT	813	

RESULT 7
US-08-018-977C-4/c
; Sequence 4, Application US/08018977C
; Patent No. 5686601
; GENERAL INFORMATION:
; APPLICANT: Weber, Peter C.
; TITLE OF INVENTION: DNA Sequences Encoding Mutant Antiviral
; TITLE OF INVENTION: Regulatory Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Pennsylvania State University
; STREET: 113 Technology Center
; CITY: University Park
; STATE: PA
; COUNTRY: US
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS 6.22
; SOFTWARE: Microsoft word for Windows 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/018,977C
; FILING DATE: 18-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/726,071
; FILING DATE: 05-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas J. Monahan
; REGISTRATION NUMBER: 29,855
; REFERENCE/DOCKET NUMBER: 91-1039/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: {814} 865-6277
; TELEFAX: {814} 865-3591
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 bases
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-018-977C-4

	Query Match	8.1%	Score 37.6	DB 1	Length 936
	Best Local Similarity	54.3%	Pred. No. 0.58		
	Matches 76	Conservative 0	Mismatches 64	Indels 0	Gaps
Qy	32	CTCGGGCGCAGCCCGGAGCCAGGCTTCCGGCGCTCTTAGCGCGCGGCGCGCGGG 91			
Db	785	CTCGGAGCGACCGCGCGCTCTGTCTCAGAGAGCGCGGAGGGAGCGGGCGGGCGCG 726			
Qy	92	GGGAGGGGAGCGAGACCGGCGACCTTAAGACACTCTCTACCTCCACGCCCAACCCCA 151			
Db	725	GAGGGGGCGCGCGGGAGGGCGCTGTGCCACCTTCACGCGCGCGCCCGCCCGGAGCCG 566			
Qy	152	CCACCTCCCGCCCACTCC 171			
Db	665	CGGCGACCGTCGACGCGC 646			


```
; APPLICANT: Daggett, Lorrie
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; TITLE OF INVENTION: HUMAN MATABOTROPIC GLUTAMATE RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Bruggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90719
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,574
; FILING DATE: 19930604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7760
; TELEFAX: 213-489-4210
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4181 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 370..4008
; OTHER INFORMATION: /product= "HUMAN MGLUR5B"
; OTHER INFORMATION: /note= "Variant of MGLUR5A with 96 base pair
; OTHER INFORMATION: Insertion between nucleotides 2599 and 3000."
US-08-072-574-9

Query Match 7.7%; Score 35.8; DB 1; Length 4181;
Best Local Similarity 50.3%; Pred. No. 2.2; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 87;

QY 40 CAGCCCGGAGCCCGAGCGCTTCCCGCGCTTTAGGCGCGGCGCGCGCGGCGGAGG 99
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3365 CAGCGCTTGGGCGCGCGCTCTGGGACTCGGGCGCGCGCGCGCGCGCGCGCG 3310
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 GAGCGACAGCGGACCGCTTACGACCTCTGACCTCCACCCCGACCCACCCACCTC 159
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3309 CGACCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 CCCCACACTCCCTAGATGTCTGCTGGGCGGCTGAAGCTCCCGCTTTAAGGGCG 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3249 GCACAGCTCTCCGCTCTGTGGGAGGCGCTTCATGACGCGCGCTTGGTGGG 3195
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-08-061-376-4/C
; Sequence 4, Application US/08061376
; Patent No. 6175000
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Djabali, Malek
; APPLICANT: Sellier, Licia
; APPLICANT: Parry, Pauline
; TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
; TITLE OF INVENTION: TRANSLOCATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIAS
```

```
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Bruggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/061,376
; FILING DATE: 13-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)546-4737
; TELEFAX: (619)546-9392
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11907 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-061-376-4

Query Match 7.7%; Score 35.8; DB 4; Length 11907;
Best Local Similarity 50.9%; Pred. No. 2.6; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 82;

QY 35 CGCGCGAGCCCGGAGCCCGAGCGCTTCCCGCGCTTTAGGCGCGGCGCGCGCGG 94
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 CGCCACAGCGCGGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 124
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 AAGGGAGCGACACCGCGGACCGCTTAAGACACCTGCTGTACCTCCACCCCGACCC 154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 GGAAGACAGCAGCGCGCGGAGCGCTTGGCGCGGCTGCGCCCTAGCGCGCGCG 64
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 ACCTCCCGCACTCCCTAGATGCTGTGGGCGGCTGAACGTCGCC 201
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 17
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```



```
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-08251-19

Query Match      7.6%; Score 35.2; DB 5; Length 35100;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY  57 CTTCCTCCGCGCTCTTAGCGCGCGCGCGCGCGGGAAGGGACGCGAGACCGCGGAGCC 116
DB  8471 COTGCCCGACACCTGTGTGAGTCTCTGCCCCCATAGATGATCGTGGAAACCATTTGAGGGGAG 8412

QY  117 CTAAGACACCTGCTGTACCTCCACCCGACCCACCCACCTCCGCCCAACTCCCTAGAT 176
DB  8411 AAGTGAAGAAAGAGGACATCCACCTGCCCCCATCAGACCCAGTCCCCAGTCCACACCC 8352

QY  177 GTGTCTGGGCGGTGAAGCTGCGCGTTAAGGGGGGGGGCGCGCTCCAGCTGC 232
DB  8351 ATGTGGCGCGGCTCTGTGCTCCGACCTTCAGGCTGGCTGGCGGTCCCATGC 8296
```

Search completed: February 3, 2003, 21:01:05
Job time : 7799.73 secs

US-09-815-048-3

Query Match 8.1%; Score 37.4; DB 10; Length 42450;
Best Local Similarity 51.5%; Pred. No. 0.95;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 5 GGTCTCCGGGACGTGGACTAGAGAGTCTGGCGGCAGCCCGGAGCCCGGCTTCGGCC 64
DB 1537 GGCTCCCGGACGCAAGCTCTAGAGGCGCGGAGAGGCGCCCGCGCCCTTCGGCC 1596
QY 65 GGGTCTTAAGCGCGGCGCGCGGCGGGAAGGAGGAGCGAGACCGCGGACCGCTTAAGACA 124
DB 1597 CCACCCACAGCCCGCGCCCGCCAGCCAGCCAGGCGCCCGCCCGCCCGCCCGCC 1656
QY 125 CCGTGTGACCTCCACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 171
DB 1657 CCAGCCCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1703

RESULT 5
US-10-047-542-44
; Sequence 44, Application US/0047542
; Patent No. US20020168367A1
; GENERAL INFORMATION:
; APPLICANT: LARRICK, JAMES W.
; APPLICANT: WYCOFF, KEITH L.
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
; FILE OF INVENTION: AND BACTERIAL DISEASES
; FILE REFERENCE: 030905.0004.CIP1
; CURRENT APPLICATION NUMBER: US/10/047,542
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US01/3932
; PRIOR FILING DATE: 2001-04-28
; PRIOR APPLICATION NUMBER: 60/200,298
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 44
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-047-542-44

Query Match 8.0%; Score 37.2; DB 9; Length 1920;
Best Local Similarity 49.5%; Pred. No. 0.5;
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 47 GAGCCCGACGCTTCCTCCGCGCTTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 106
DB 508 GAACACCCAGCATGCTGGGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 567
QY 107 ACGGCGGACCTTAAGACACCTGCTGTACCTCCACCCCGCCCGCCCGCCCGCCCGCC 156
DB 568 AGCTCAGAGCGCAACACCTCTATGACACACAGCTCTCCCGCCCGCGCGCGCGCGCG 627
QY 167 CTCCTCPAGATGTCTGCGCGCGCTGAAGCTCCCGCTTTAAGGGCGCGCGCGCGCGCG 226
DB 628 TTCACCCCGCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 687
QY 227 ACCTGCTTCTGCT 240
DB 688 CCGACCAATCCAGT 701

RESULT 6
US-09-799-462A-17
; Sequence 17, Application US/09799462A
; Patent No. US20020160970A1
; GENERAL INFORMATION:
; APPLICANT: HADIACZKY, GYULA
; APPLICANT: SZALAY, ALADAR
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
; AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,462A
FILING DATE: 10-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,682
FILING DATE: 10-APR-1997
APPLICATION NUMBER: 08/695,191
FILING DATE: 07-AUG-1996
APPLICATION NUMBER: 08/682,080
FILING DATE: 15-JUL-1996
APPLICATION NUMBER: 08/629,822
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24601-402G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8403
TELEFAX: 858-587-5360
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 42999 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: <Unknown>
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-799-462A-17

Query Match 8.0%; Score 37.2; DB 9; Length 42999;
Best Local Similarity 51.9%; Pred. No. 1.1;
Matches 84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 14 GACGTGGACTAGAGAGTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
DB 10013 GAGAGTGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10072
QY 74 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
DB 10073 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10132
QY 134 CCTCCACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 175
DB 10133 TCGTCCCTCTCTCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10174

RESULT 7
US-10-125-767-17
; Sequence 17, Application US/10125767
; Patent No. US20020160410A1
; GENERAL INFORMATION:
; APPLICANT: HADIACZKY, GYULA
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
; METHODS

```
; FOR PREPARING ARTIFICIAL CHROMOSOMES
;
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Erman White & McAuliffe LLP
; STREET: 4350 La Jolla Village Drive, 7th Floor
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92122
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
;
; CURRENT APPLICATION DATA: US/10/125,767
; FILING DATE: 17-Apr-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/724,693
; FILING DATE: 28-NOV-2000
; APPLICATION NUMBER: 08/835,682
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: 08/695,191
; FILING DATE: 07-AUG-1996
; APPLICATION NUMBER: 08/682,080
; FILING DATE: 15-JUL-1995
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24601-402J
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-450-8403
; TELEFAX: 858-587-5360
;
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-125-767-17

Query Match 8.0%; Score 37.2; DB 9; Length 42999;
Best Local Similarity 51.9%; Pred. No. 1.1;
Matches 84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 14 GAGGTGAGTACAGATCTGCGCGGAGCCGAGCCGAGCCGCTCCCGGGGCTTTAG 73
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10013 GAGAGTGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10072

QY 74 GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10073 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10132

QY 134 CCCTCCACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 175
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10133 TCCTCCCTCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10174

RESULT 8
US-09-772-304A-1/c
; Sequence 1, Application US/09772304A
; Patent No. US20020146785A1
; GENERAL INFORMATION:
; APPLICANT: Manishi, L.H.
; TITLE OF INVENTION: NOVEL ESCHERICHIA COLI HAVING ACCESSION
; FILE REFERENCE: A33943 066123.0103
; CURRENT APPLICATION NUMBER: US/09/772,304A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4826
; TYPE: DNA
; ORGANISM: Streptomyces aureofaciens
US-09-772-304A-1

Query Match 7.9%; Score 36.6; DB 10; Length 4826;
Best Local Similarity 49.7%; Pred. No. 0.91;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 42 GCGCCGAGCCGCGCTTCCGCGCGCTTTAGCGCGCGCGCGCGCGCGCGCGCGGA 101
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 947 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 102 CGCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 161
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 887 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 828
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 CCAACTCCCTAGATGTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 221
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 827 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 768
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 GCTCCAC 228
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 767 CCGCGCGCG 761
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-833-381-1386
; Sequence 1386, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1386
; LENGTH: 1102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1102)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1386

Query Match 7.8%; Score 36.2; DB 10; Length 1102;
Best Local Similarity 51.2%; Pred. No. 0.8;
Matches 109; Conservative 0; Mismatches 103; Indels 1; Gaps 1;

QY 123 CACTGCTGTACCTCCACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 CACTCCAGTCTGTGCTGCTTACCTCCAGCCACCATCTGGCGCACTGCTAGTAGCG 724
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 TGGCGCGCTGACGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 TCAGGGGATCCAGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 783
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
QY 243 GTGACTGAACATACAAACAGAGAGCGCGGGGAGGAGGAGGAGACAGCGT 302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 CAGCATGATCGCTCTCTGTCAGTGCAGGGGCGCGGAGCGGTAGCGCCCGGGCC 843
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 303 TTGACCGATAGTAACCTCTGGCTCGGTGCACG 335
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 GGGCTGGCGGCGAGCGCGCGCTCGGGAGGC 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-834-975-987/c
; Sequence 987, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 987
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1797)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-987

Query Match 7.7%; Score 35.6; DB 10; Length 1797;
Best Local Similarity 49.0%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 0;

QY 35 CGGCGAGCGCCGAGCCAGCGCTTCCCGCGGCTTAGCGCGCGCGCGCGCGGGG 94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1585 CGGGCGCTGCCACCGCACCCACCGCGCGCGCGCGCGCGCGGAGGAGGCGG 1526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 AAGGGAGCGAGACCGGGGACCTTAAGACACCTGCTGTACCCICACCCCGCCACCC 154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1525 CAGGAAGCGCGGGGCTCTCCCTCCGCGTCCCGGGCGCGCGAGTTCCGGAAGCCCTCC 1466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 ACCTCCCGCAACTCCCTAGATGTGTGCGGGGCTGAACGTGCGCCGTTAAGGGCG 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1465 CGCTCCTCCCTGGCGGAGCGCGCGCGCGCGGGAATGGCGGATGAGCCTCC 1406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 GCGCGCGCTCCAC 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 GGAGCGCGCTCCCC 1392
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-834-975-1003/c
; Sequence 1003, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
```

```
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1003
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1797)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-1003

Query Match 7.7%; Score 35.6; DB 10; Length 1797;
Best Local Similarity 49.0%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 0;

QY 35 CGGCGAGCGCCGAGCCAGCGCTTCCCGCGGCTTAGCGCGCGCGCGCGGGG 94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1585 CGGGCGCTGCCACCGCACCCACCGCGCGCGCGCGCGCGGAGGAGGCGG 1526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 AAGGGAGCGAGACCGGGGACCTTAAGACACCTGCTGTACCCICACCCCGCCACCC 154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1525 CAGGAAGCGCGGGGCTCTCCCTCCGCGTCCCGGGCGCGCGAGTTCCGGAAGCCCTCC 1466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 ACCTCCCGCAACTCCCTAGATGTGTGCGGGGCTGAACGTGCGCCGTTAAGGGCG 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1465 CGCTCCTCCCTGGCGGAGCGCGCGCGCGCGGGAATGGCGGATGAGCCTCC 1406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 GCGCGCGCTCCAC 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 GGAGCGCGCTCCCC 1392
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-834-975-1012/c
; Sequence 1012, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF HUMAN CANCERS
; FILE REFERENCE: MRI-0163
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1012
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1797)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-1012

Query Match 7.7%; Score 35.6; DB 10; Length 1797;
Best Local Similarity 49.0%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 0;

QY 35 CGGCGAGCGCCGAGCCAGCGCTTCCCGCGGCTTAGCGCGCGCGCGGGG 94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1585 CGGGCGCTGCCACCGCACCCACCGCGCGCGCGCGCGGAGGAGGCGG 1526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 AAGGGAGCGAGACCGGGGACCTTAAGACACCTGCTGTACCCICACCCCGCCACCC 154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1525 CAGGAAGCGCGGGGCTCTCCCTCCGCGTCCCGGGCGCGCGAGTTCCGGAAGCCCTCC 1466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 ACCTCCCGCAACTCCCTAGATGTGTGCGGGGCTGAACGTGCGCCGTTAAGGGCG 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1465 CGCTCCTCCCTGGCGGAGCGCGCGCGCGCGGGAATGGCGGATGAGCCTCC 1406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 GCGCGCGCTCCAC 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 GGAGCGCGCTCCCC 1392
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-834-975-1013/c
; Sequence 1013, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF HUMAN CANCERS
; FILE REFERENCE: MRI-0164
; CURRENT APPLICATION NUMBER: US/09/834,975
```


Tue Feb 4 09:39:45 2003

us-09-596-141c-3_copy_1181_1643.rnpb

Search completed: February 3, 2003, 16:31:11
Job time : 95.1333 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:39:45 ; Search time 666.044 Seconds
(without alignments)
8559.210 Million cell updates/sec

Title: us-09-596-141c-3_copy_1292_1643
Perfect score: 352
Sequence: 1 ggaacctaaacacactgtgtg.....cagggcagggcgaggagctc 352

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32306132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estm: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_estl: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89.4	25.4	736	9 AUI35588	AUI35588 AUI35588
2	89	25.3	547	9 AL698654	AL698654 DKF23686N
3	89	25.3	763	9 AUI21731	AUI21731 AUI21731
4	75	21.3	292	14 Z44377	Z44377 HSC12B081 n
5	75	21.3	1004	11 BC034824	BC034824 Homo sapi
6	60.2	17.1	998	12 BG678861	BG678861 602624760

C 7	47.6	13.5	925	17	CNS0091P	AL053013 Drosophil
C 8	43	12.2	935	17	CNS006XK	AL066051 Drosophil
C 9	42.4	12.0	910	17	CNS006GN	AL065629 Drosophil
C 10	42.2	12.0	925	17	CNS0091P	AL053013 Drosophil
C 11	39.6	11.2	912	17	CNS006N3	AL065775 Drosophil
C 12	39.4	11.2	1101	17	CNS00397	AL063912 Drosophil
C 13	38.8	11.0	384	17	CNS035G7	AL228688 Tetradon
C 14	38.8	11.0	827	17	CNS0137T	AL102515 Drosophil
C 15	38.8	11.0	844	17	CNS0052P	AL056652 Drosophil
C 16	38.8	11.0	885	13	BM415333	BM415333 Op20407 M
C 17	38.4	10.9	490	9	AL580352	AL580352 A580352
C 18	38.4	10.9	1101	17	CNS017Sx	AL108460 Drosophil
C 19	38.4	10.9	1257	14	BQ675446	BQ675446 AGENCOURT
C 20	37.8	10.7	908	13	BM416197	BM416197 Op21284 M
C 21	37.8	10.7	1863	17	GGA200110	AJ231825 Gallus ga
C 22	37.2	10.6	1177	14	BQ677491	BQ677491 AGENCOURT
C 23	37.2	10.6	1201	17	CNS016BR	AL106545 Drosophil
C 24	36.8	10.5	840	12	BF265554	BF265554 HV_GEA001
C 25	36.8	10.5	1101	17	CNS012JV	AL101653 Drosophil
C 26	36.6	10.4	881	17	CNS00512	AL065702 Drosophil
C 27	36.6	10.4	1203	17	CNS015V4	AL106054 Drosophil
C 28	36.4	10.3	1201	17	CNS007K1	AL067365 Drosophil
C 29	36.2	10.3	833	14	BQ219067	BQ219067 AGENCOURT
C 30	36.2	10.3	933	17	AG073525	AG073525 Pan trogl
C 31	36	10.2	384	17	CNS043G6	AL272751 Tetradon
C 32	36	10.2	740	14	BQ178338	BQ178338 Ut-M-ERO-
C 33	36	10.2	966	17	CNS06SUF	AL413725 17 end of
C 34	36	10.2	1101	17	CNS00GP3	AL072367 Drosophil
C 35	36	10.2	1409	12	BF303894	BF303894 60186514
C 36	36	10.2	1446	17	AG072447	AG072447 Pan trogl
C 37	35.8	10.2	487	17	CNS0060P	AL062040 Drosophil
C 38	35.8	10.2	1011	17	CNS011KG	AL316969 Tetradon
C 39	35.8	10.2	1100	14	BM921360	BM921360 AGENCOURT
C 40	35.8	10.2	1101	17	CNS00EPO	AL069493 Drosophil
C 41	35.8	10.2	1201	17	CNS015VQ	AL106076 Drosophil
C 42	35.6	10.1	513	10	BES18114	BE618114 601462521
C 43	35.6	10.1	1101	17	CNS01720	AL107514 Drosophil
C 44	35.4	10.1	809	17	AG136883	AG136883 Pan trogl
C 45	35.4	10.1	857	17	BH157466	BH157466 ENTP347F

ALIGNMENTS

RESULT 1
AUI35588
LOCUS AUI35588 PLACE1 Homo sapiens CDNA linear EST 02-AUG-2002
DEFINITION AUI35588 PLACE1 Homo sapiens CDNA c1cne PLACE1002437 5', mRNA
ACCESSION AUI35588
VERSION AUI35588.1 GI:10996127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 736)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,C., Wakamatsu,A., Nakamura,Y., Nagei,T., Sugano,S. and Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0612, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

```

FEATURES
  source
    Location/Qualifiers
      1. .736
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="PLACE:002437"
        /clone_lib="2ACEL"
        /tissue_type="placenta"
        /note="vector: pME18SFL3"
BASE COUNT      163 a   199 c   199 g   170 t      5 others
ORIGIN

Query Match      25.4%; Score 89.4; DB 9; Length 736;
Best Local Similarity 98.9%; Pred. No. 1.6e-13;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 262 GTAATTCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTC 321
|||||
DB 1 GTAGTGGGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTC 60
|||||

QY 322 CCGGCTCGGCGAGGCGAGGCGGGGAGCTC 352
|||||
DB 61 CCGGCTCGGCGAGGCGAGGCGGGGAGCTC 91
|||||

RESULT 2
AL698654
LOCUS      AL698654      547 bp      mRNA      linear      EST 21-MAR-2002
DEFINITION DKFZp686N12109_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION  AL598654
VERSION    AL598654
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 547)
AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
          S.
          EST (Duesterhoeft, et al.)
          Unpublished (1999)
          Contact: Duesterhoeft A
          MIPS
          Am Klopferspitz 18a D-82152 Martinsried, Germany
          This is the 5' sequence of the clone insert
          Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
          Research Center (DKFZ); Email: s.wiemann@kfz-heidelberg.de;
          sequenced by Qiagen (Hilden/Germany); within the cDNA sequencing
          consortium of the German Genome Project.
          No si sequence available.
          This clone (DKFZp686N12109) is available at the RZPD in Berlin.
          Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
          Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
          Location/Qualifiers
            1. .547
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="DKFZp686N12109"
              /clone_lib="586 (synonym: hlcc3)"
              /dev_stage="adult"
              /tissue_type="human skeletal muscle"
              /lab_host="DH102"
              /note="vector: pTriplex2; Site_1: SfIIA; Site_2: SfiIB;
              cDNA-collection"
BASE COUNT      112 a   154 c   157 g   124 t
ORIGIN

Query Match      25.3%; Score 89; DB 9; Length 547;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 AATTCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTCC 323
|||||

```

```

Db 2 AATTGGAGCGAGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTCC 61
QY 324 GGGCTGGGCGAGGCGAGGCGGGGAGCTC 352
|||||
Db 62 GGGCTGGGCGAGGCGAGGCGGGGAGCTC 90
|||||

RESULT 3
AU121731
LOCUS      AU121731      763 bp      mRNA      linear      EST 01-AUG-2002
DEFINITION AU121731 MAMMAL Homo sapiens cDNA clone MAMMAL000851 5', mRNA
sequence.
ACCESSION  AU121731
VERSION    AU121731.1 GI:10936966
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 763)
AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
          Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
          isogai,T.
          HRI human cDNA project
          Unpublished (2000)
          Contact: Takao Isogai
          Genomics Laboratory
          Heix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: genomics@hri.co.jp
          HRI human cDNA project; 5' - & 3'-end one pass sequencing: Helix
          Research Institute; cDNA library construction: Department of
          Virology, Institute of Medical Science, University of Tokyo, and
          Heix Research Institute.
          Location/Qualifiers
            1. .763
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="MAMMAL000851"
              /clone_lib="MAMMAL"
              /tissue_type="mammary gland"
              /note="Vector: pME18SFL3"
BASE COUNT      137 a   205 c   260 g   158 t      3 others
ORIGIN

Query Match      25.3%; Score 89; DB 9; Length 763;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 AATTGGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTCC 323
|||||
Db 1 AATTGGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGCGAGCCCTTCTCTCC 60
|||||

QY 324 GGGCTGGGCGAGGCGAGGCGGGGAGCTC 352
|||||
Db 61 GGGCTGGGCGAGGCGAGGCGGGGAGCTC 89
|||||

RESULT 4
Z44377
LOCUS      Z44377      292 bp      mRNA      linear      EST 14-NOV-1994
DEFINITION HSC125081 normalized infant brain cDNA Homo sapiens cDNA clone
c-12b08, mRNA sequence.
ACCESSION  Z44377
VERSION    Z44377.1 GI:573506
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 292)

```

AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Boulgatte, R., Jumelet, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y., Senastiani-Kabaktchis, C. and Tessier, A.

TITLE IMAGE: molecular integration of the analysis of the human genome and its expression

JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

MEDLINE 95277534

COMMENT Contact: Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel.: 33169472800
Fax: 33160778698
Email: genethon@genethon.fr

FEATURES Source
1..292
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-lzb08"
/clone_lib="normalized infant brain cdna"
/sex="Female"
/tissue.type="total brain"
/dev_stage="3 months old"
/note="Origin: Brain; Vector: lambdafmid BA; Size: 1.1 kb; HindIII; Site 2: NotI; sex: Female; dev_stage=3 months old; isolate: muscular atrophy patient; tissue.type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lambdafmid BA vector. Clone library from B. Soares, Psychiatry Dept. Columbia University, USA. Normalization method: Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 96 g 56 t

ORIGIN

Query Match 21.3%; Score 75; DB 14; Length 292;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 GTGAGTGGGGCGGGACCGGAGCGAGCGCGACCGCTTCTCTCCCGGGTGCGGAGGG 337
|||||
Db 1 GTGAGTGGGGCGGGACCGGAGCGAGCGCGACCGCTTCTCTCCCGGGTGCGGAGGG 60

QY 338 CAGGGCGGGAGCTC 352
|||||
Db 61 CAGGGCGGGAGCTC 75

RESULT 5
LOCUS BC034824
DEFINITION Homo sapiens, clone IMAGE:4749735, mRNA.
ACCESSION BC034824
VERSION BC034824.1 GI:21961568
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1004)
Strasbourg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarsson, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 52 Row: d Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5915657
This clone has the following problem: incomplete processing.

FEATURES Location/Qualifiers
1..1004
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749735"
/tissue.type="Skin, squamous cell carcinoma"
/clone_lib="NCI CGAP_Skn4"
/lab_host="DH10B"
/note="Vector: PCWV-SPORT6.ccdB"

BASE COUNT 306 a 234 c 230 g 234 t

ORIGIN

Query Match 21.3%; Score 75; DB 11; Length 1004;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 GTGAGTGGGGCGGGACCGGAGCGCGAGCGCGACCGCTTCTCTCCCGGGTGCGGAGGG 337
|||||
Db 1 GTGAGTGGGGCGGGACCGGAGCGCGAGCGCGACCGCTTCTCTCCCGGGTGCGGAGGG 60

QY 338 CAGGGCGGGAGCTC 352
|||||
Db 61 CAGGGCGGGAGCTC 75

RESULT 6
LOCUS BC678861
DEFINITION mRNA sequence.
ACCESSION BC678861
VERSION BC678861.1 GI:13910258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
plate: LLAM10603 row: g column: 16
High quality sequence stop: 860.

FEATURES Location/Qualifiers
1..998
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/clone="IMAGE:4749735"
/clone_lib="NCL_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT      285 a   233 c   244 g   236 t
ORIG:N

Query Match      17.1%; Score 60.2; DB 12; Length 998;
Best Local Similarity 89.0%; Pred. No. 9.8e-06;
Matches 65; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 GAGTGGGCGGAGCCGAGCGAGCGGACCGACCTTCTCCCGGGGTGCGGCGAGGCA 339
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 GTGAGTGGGCGGAGCGGAGCGGAGCGGACCGACCTTCTCCCGGGGTGCGGCGAGGCA 60

QY 340 GCGGCGGAGCTC 352
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 GCGGCGGAGCTC 73

RESULT 7
CNS0091P/C
LOCUS      CNS0091P      925 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BAC191916 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL053013
VERSION     AL053013.1 GI:4934461
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster.
ORGANISM   Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 925)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osogawa and
            Aaron Mammose in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
            Location/Qualifiers
                1..925
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="5ACR19D16"
                /clone_lib="RPCI-98"
                /note="end : TET3"
BASE COUNT      120 a   61 c   51 g   172 t   511 others
ORIGIN

Query Match      13.5%; Score 47.6; DB 17; Length 925;
Best Local Similarity 12.8%; Pred. No. 0.022; Indels 0; Gaps 0;
Matches 42; Conservative 152; Mismatches 134; Indels 0; Gaps 0;

```

```

QY 23 CCTCCACCCACCCACCCACCCACCTCCGCCCACTCCCTAGATGTGTGTCGGCGGCTCA 82
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 935 CSBSCSCSCSBSBSCSSSSSTSSNSBSCSSSSBSSBSSSTSSMSSSSBSSSSGSSS 866
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 83 ACGTGCGCGGTAAAGGGCGGCGGCGGCTCCAGCTCTTCTCTCTGAGTGACTGAAT 142
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 865 SSGTSSACVKCNASSSCCGCGGMAACCCMCSSSSSCCGSASARGVKVRSAGGAKRGGS 806
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 143 ACATAAACAGAGCGCGGAACGGGCGGAGGAGGAGACAGCAGGCTTGACACGATAG 202
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 805 GCASASHSSSSACSSSSSSSCASCSWSSSSSSASSSSSRSGGAGGSSASSSSSSS 746
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 203 TAACCTCTGCGTCTGTCAGCGCACTATATAAAGGAAGTACTGCCGGCAAAACCCCG 262
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 745 ASAGSVSSASSSSSSCSSVSCSVASVSMSCSSBSSSSASASSSSSSKASCSCCC 686
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 263 TAATTCGAGCGAGAGTGAAGTGGCGCGGAGCCCGAGCGGAGCGGAGCGGCTCTCTCC 322
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 685 TSWSCSCSTASMSAARSSSSSSSSSSSSSASSSSSASSSSSSSSSSSSSSSSACGB 626
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 323 CGGCGTGGCGGAGGCGGCGGCGGAGC 350
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 625 SMSGCGGSGSVASSGMSVSSSGGGS 598
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
CNS006XK/C
LOCUS      CNS006XK      935 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL066051
VERSION     AL066051.1 GI:4245019
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster.
ORGANISM   Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 935)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osogawa and
            Aaron Mammose in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
            Location/Qualifiers
                1..935
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="BACR14N09"
                /clone_lib="RPCI-98"
                /note="end : T7"
BASE COUNT      257 a   170 c   162 g   96 t   250 others
ORIGIN

Query Match      12.2%; Score 43; DB 17; Length 935;
Best Local Similarity 29.8%; Pred. No. 0.37; Indels 0; Gaps 0;
Matches 95; Conservative 69; Mismatches 155; Indels 0; Gaps 0;

```


Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6&st2=RC6-BT0252-271099-012-cl0&st3=999-10-27&st4=1)
Seq primer: puc 18 forward
High quality sequence start: 146
High quality sequence stop: 289.
Location/Qualifiers
1..289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0252"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site: 1; SmaI; Site: 2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
94 a 70 c 51 g 74 t

FEATURES
source
BASE COUNT
ORIGIN
Query Match 14.9%; Score 245.4; DB 10; Length 289;
Best Local Similarity 96.9%; Pred. No. 5.3e-39;
Matches 282; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
QY 678 GTCGACITTCACAAATGATACAACTAATACAACTCTGCTGTTTATCAGGAGG 737
Db 289 GTCGACITTCACAAATGATACAACTAATACAACTCTGCTGTTTATCAGGAGG 230
QY 738 CTGATCAATATATGAATATAAAGGGGCTGCTCCATATGTTCTGTTTCTTTG 797
Db 229 CTGATCAATATATGAATATAAAGGGGCTGCT -CCATATTCTGTTCTTTGTT 171
QY 798 TTGTTTCTTTTCTTTTCTTTGCTGCTCTCTCTCAATATATGAAGAGACAGTAA 857
Db 170 TTGTTTCTTTTCTTTTCTTTGCTGCTCTCTCTCAATATATGAAGAGACAGTAA 112
QY 858 GATGTTCTCTCGGGTCTCTGAGGAGACCTGAGGAGTCAAGTGGGAATCTCCAGGCA 917
Db 111 GATGTTCTCTCGGGTCTCTGAGGAGACCTGAGGAGTCAAGTGGG -ATCTCCAGGCA 53
QY 918 GTAGTGGCTATCAAAATCAAAATCAAGTCCAGTTTGTGGGGGAAACAAA 968
Db 52 GTAGTGGCTATCAAAATCAAAATCAAGTCCAGTTTGTGGGGGAAACAAA 2

RESULT 2
AA527406
LOCUS
DEFINITION
mRNA sequence.
AA527406
VERSION
AA527406.1 GI:2269475
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

FEATURES
source
BASE COUNT
ORIGIN
Query Match 14.9%; Score 245.4; DB 10; Length 289;
Best Local Similarity 96.9%; Pred. No. 5.3e-39;
Matches 282; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
QY 678 GTCGACITTCACAAATGATACAACTAATACAACTCTGCTGTTTATCAGGAGG 737
Db 289 GTCGACITTCACAAATGATACAACTAATACAACTCTGCTGTTTATCAGGAGG 230
QY 738 CTGATCAATATATGAATATAAAGGGGCTGCTCCATATGTTCTGTTTCTTTG 797
Db 229 CTGATCAATATATGAATATAAAGGGGCTGCT -CCATATTCTGTTCTTTGTT 171
QY 798 TTGTTTCTTTTCTTTTCTTTGCTGCTCTCTCTCAATATATGAAGAGACAGTAA 857
Db 170 TTGTTTCTTTTCTTTTCTTTGCTGCTCTCTCTCAATATATGAAGAGACAGTAA 112
QY 858 GATGTTCTCTCGGGTCTCTGAGGAGACCTGAGGAGTCAAGTGGGAATCTCCAGGCA 917
Db 111 GATGTTCTCTCGGGTCTCTGAGGAGACCTGAGGAGTCAAGTGGG -ATCTCCAGGCA 53
QY 918 GTAGTGGCTATCAAAATCAAAATCAAGTCCAGTTTGTGGGGGAAACAAA 968
Db 52 GTAGTGGCTATCAAAATCAAAATCAAGTCCAGTTTGTGGGGGAAACAAA 2

RESULT 2
AA527406
LOCUS
DEFINITION
mRNA sequence.
AA527406
VERSION
AA527406.1 GI:2269475
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1057 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 380.
Location/Qualifiers
1..467
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CCGAP Co3"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site: 1; Not I; Site: 2; Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization."
72 a 118 c 166 g 111 t

FEATURES
source
BASE COUNT
ORIGIN
Query Match 6.8%; Score 111.4; DB 9; Length 467;
Best Local Similarity 99.1%; Pred. No. 3e-12;
Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 816 TTGTGGCTCTCTCTCTCAATTTATGAAGAGACAGTAAAGTCTCTCTCGGTC 875
Db 1 TTGTGGCTCTCTCTCTCAATTTATGAAGAGACAGTAAAGTCTCTCTCGGTC 60
QY 876 TGTAGGACCTGGGAGGAGCTCAGGCTGGGAAATCTCCAGGAGAGTGGCCT 928
Db 61 TGTAGGACCTGGGAGGAGCTCAGGCTGGGAAATCTCCAGGAGAGTGGCCT 113

RESULT 3
AU135588
LOCUS
DEFINITION
sequence.
AU135588
VERSION
AU135588.1 GI:10996127
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 736)
Ota, T., Nishikawa, I., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5' & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
Location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"

FEATURES
source
BASE COUNT
ORIGIN
Query Match 6.8%; Score 111.4; DB 9; Length 467;
Best Local Similarity 99.1%; Pred. No. 3e-12;
Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 816 TTGTGGCTCTCTCTCTCAATTTATGAAGAGACAGTAAAGTCTCTCTCGGTC 875
Db 1 TTGTGGCTCTCTCTCTCAATTTATGAAGAGACAGTAAAGTCTCTCTCGGTC 60
QY 876 TGTAGGACCTGGGAGGAGCTCAGGCTGGGAAATCTCCAGGAGAGTGGCCT 928
Db 61 TGTAGGACCTGGGAGGAGCTCAGGCTGGGAAATCTCCAGGAGAGTGGCCT 113

RESULT 3
AU135588
LOCUS
DEFINITION
sequence.
AU135588
VERSION
AU135588.1 GI:10996127
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 736)
Ota, T., Nishikawa, I., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5' & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
Location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1002437"

<hr/>					
RESULT 5					
AUI21731		763 bp	mRNA	linear	EST 01-AUG-2002
<hr/>					
LOCUS	AUI21731	MAMMAL	Homo sapiens cDNA clone	MAMMAL000851	5', mRNA
<hr/>					
DEFINITION	AUI21731	MAMMAL	Homo sapiens cDNA clone	MAMMAL000851	5', mRNA
<hr/>					
sequence.					
ACCESSION	AUI21731				
VERSION	AUI21731				
KEYWORDS	EST,				
SOURCE	AUI21731.1	GI:10936966			
<hr/>					
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE	1	(bases 1 to 763)			
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yanamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.				
<hr/>					
TITLE	HRI human cDNA project				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isoqai Genomics Laboratory Helix Research Institute 1532-3 Yara, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5'- s 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..763 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MAMMAL000851" /clone_lib="MAMMAL" /tissue_type="mammary gland" /note="Vector: pME18SPL3"				
<hr/>					
FEATURES					
source	137 a	205 c	260 g	158 t	3 others
<hr/>					
BASE COUNT					
ORIGIN					
Query Match	5.4%	Score 89;	DB 9;	Length 763;	
Best Local Similarity	100.0%;	Pred.No. 7.5e-08;			
Matches 89;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
<hr/>					
QY	1555	AATTCGGAGCAGAGTGTGCTGGCGGCAGCCGAGCGCGACGCCCTCTCTCC	1614		
Dd	1	AATTCGGAGCAGAGTGTGCTGGCGGCAGCCGAGCGCGACGCCCTCTCTCC	60		
<hr/>					
QY	1615	GGGCTGGCGAGCGGCGAGCGGGGAGCTC	1643		
Dd	61	GGGCTGGCGAGCGGCGAGCGGGGAGCTC	89		
<hr/>					
RESULI	6				
LOCUS	Z44377				
DEFINITION	HSCIZB081	normalized infant brain cDNA	Homo sapiens cDNA clone		EST 14-NOV-1999
c-izb08, mRNA sequence.					
ACCESSION	Z44377				
VERSION	Z44377.1	GI:573506			
KEYWORDS	EST,				
SOURCE	human.				
<hr/>					
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE	1	(bases 1 to 292)			
AUTHORS	Aufrey,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houllgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,E., Marriage-Samsom,R., Pietu,G., Pouliot,Y., Sebastiani-Kabatchis,C. and Tessier.A.				
<hr/>					
TITLE	IMAGE: molecular integration of the analysis of the human genome				

TITLE The zinc finger protein 202 (ZNF202) is a transcriptional repressor of AIP binding cassette transporter A1 (ABCA1) and ABCG1 gene expression and a modulator of cellular lipid efflux

JOURNAL J. Biol. Chem. 276 (15), 12427-12433 (2001)

J. Biol. Chem. 276 (15), 12427-12433 (2001))
JOURNAL
MEDLINE
21192304
PUBMED
1-279031
REFERENCE
2 (bases 1 to 1167)
AUTHORS
Porcch-Oezcuernuez, M. K.
TITLE
Direct Submission
Submitted (05-JAN-2000) Porcch-Oezcuernuez M. K., Institute for
Clinical Chemistry, University of Regensburg,
Franz-Josef-Strauss-Allee 11 93042 Regensburg, GERMANY

```

FEATURES
    Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY
    Location/Qualifiers
        source
            1..1167
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="9"
                /map="9q22-31"
                /cell_type="leukocyte"
            1..1167
                /gene="ABC-1"
            1..942
                promote;

```

```
promoter=
1..942
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
896..900
TATA_signal
/gene="ABC-1"
intron
148..1187
/gene="ABC-1"
```

```

      /Genes:"ABC-1"
      /function="cholesterol efflux regulatory protein"
278 a      315 C      327 g      247 t
      IGIN
Query Match      100.0%      Score 139; DB 9; Length 1167.
Seq. Local Similarity      100.0%

```

Query match 100.0%; Score 239; DB 9; Length 1157;
Best Local Similarity 100.0%; Pred. NO. 1.2e-31;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Caps 0;

1 GGGCCCCGGCTCCACGTGCTTCTGCTAGTGACTGAACATAAACAGAGGCCGGAA 60
768 GGGCCCCGGCTCCACGTGCTTCTGCTAGTGACTGAACATAAACAGAGGCCGGAA 827

788 GGGCCCCGGCTCCAGGTCCTTCTGCTAGCTGACTGAACATATAACAGAGGCCGGGAA 827

61 CGGGCGGGGAGGAGGAGACACAGGCTTGACCGCTACTTAACCTCTGGCTCGTGTCA 120

828 CGGGCGGGGAGGAGGAGACACAGGCTTGACCGCTACTTAACCTCTGGCTCGTGTCA 887

121 GCCGAATCTATAAAGGAA 139

121 GCCGAATCTATAAAGGAA 139
|||||
888 GCCGAATCTATAAAGGAA 906

RESULT 2
 AX060715
 CUS
 AX060715
 DEFINITION
 Sequence 3 from Patent WO0078972.
 AX060715
 CREATION
 AX060715.1
 CREATION
 AX060715.1
 GI:12406104
 KEYWORDS

YWORDS
URGE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1643)
Lamp R M Wade D and Garrison M

REFERENCE
AUTHORS
TITLE
JOURNAL
CV THERAPEUTICS, INC. (US)
LOCATION/Qualifiers
source

```
source
1..1643
/organism="Homo sapiens"
/db_xref="taxon:9606"
370 a 413 c 457 g 403 t
--SE COUNT
--GIN
Query Match 100.0% Score 128. DB E-Value 1e-42
```

Query Match 100.0%; Score 139; DB 6; Length 1643:-

Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeAtellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Hesford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehotsky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
MacDonald, P., Marcuis, N., McEvan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gl:6705871.

COMMENT

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seg.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: L_N10

NOTE: This record contains 73 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

1
871: contig of 871 bp in length
872 971: gap of 100 bp
972 1834: contig of 863 bp in length
1835 1934: gap of 100 bp
1935 2804: contig of 870 bp in length
2805 2904: gap of 100 bp
2905 3745: contig of 841 bp in length
3746 3845: gap of 100 bp
3846 4596: contig of 851 bp in length
4597 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 6540: contig of 800 bp in length
6541 6640: gap of 100 bp
6641 7509: contig of 869 bp in length
7510 7609: gap of 100 bp
7610 8473: contig of 870 bp in length
8480 8579: gap of 100 bp
8580 9430: contig of 851 bp in length
9431 9530: gap of 100 bp
9531 10376: contig of 846 bp in length
10377 10476: gap of 100 bp
10477 11322: contig of 846 bp in length
11323 11422: gap of 100 bp
11423 12303: contig of 880 bp in length
12303 12402: gap of 100 bp
12403 13280: contig of 878 bp in length
13281 13380: gap of 100 bp
13381 14241: contig of 861 bp in length
14242 14341: gap of 100 bp

14342 15127: contig of 855 bp in length
15127 15127: gap of 100 bp
15297 16123: contig of 827 bp in length
16224 16223: gap of 100 bp
16224 17072: contig of 849 bp in length
17073 17172: gap of 100 bp
17173 18041: contig of 869 bp in length
18042 18441: gap of 100 bp
18442 19009: contig of 868 bp in length
19010 19109: gap of 100 bp
19110 19966: contig of 857 bp in length
19967 20066: gap of 100 bp
20067 20921: contig of 855 bp in length
20922 21021: gap of 100 bp
21022 21865: contig of 844 bp in length
21866 21955: gap of 100 bp
21956 22832: contig of 867 bp in length
22833 22932: gap of 100 bp
22933 23780: contig of 848 bp in length
23781 23880: gap of 100 bp
23881 24733: contig of 853 bp in length
24734 24833: gap of 100 bp
24834 25670: contig of 837 bp in length
25671 25770: gap of 100 bp
25771 26621: contig of 851 bp in length
26622 26721: gap of 100 bp
26722 27576: contig of 855 bp in length
27577 27676: gap of 100 bp
27677 28532: contig of 856 bp in length
28533 28632: gap of 100 bp
28633 29492: contig of 860 bp in length
29493 29592: gap of 100 bp
29593 30455: contig of 863 bp in length
30456 30555: gap of 100 bp
30556 31411: contig of 855 bp in length
31411 31510: gap of 100 bp
31511 32368: contig of 858 bp in length
32369 32468: gap of 100 bp
32469 33312: contig of 844 bp in length
33313 33412: gap of 100 bp
33413 34268: contig of 856 bp in length
34269 34368: gap of 100 bp
34369 35204: contig of 836 bp in length
35205 35304: gap of 100 bp
35305 36156: contig of 852 bp in length
36157 36256: gap of 100 bp
36257 37128: contig of 872 bp in length
37129 37228: gap of 100 bp
37229 38083: contig of 855 bp in length
38084 38183: gap of 100 bp
38184 39031: contig of 848 bp in length
39032 39131: gap of 100 bp
39132 40006: contig of 875 bp in length
40007 40106: gap of 100 bp
40107 40967: contig of 861 bp in length
40968 41067: gap of 100 bp
41068 41913: contig of 846 bp in length
41914 42013: gap of 100 bp
42014 42824: contig of 811 bp in length
42825 42924: gap of 100 bp
42925 43776: contig of 852 bp in length
43777 43876: gap of 100 bp
43877 44752: contig of 876 bp in length
44753 44852: gap of 100 bp
44853 45724: contig of 872 bp in length
45725 45824: gap of 100 bp
45825 46643: contig of 819 bp in length
46644 46743: gap of 100 bp
46744 47599: contig of 856 bp in length
47599 47699: gap of 100 bp
47700 48551: contig of 852 bp in length
48552 48651: gap of 100 bp
48652 49485: contig of 834 bp in length

```

* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54258: contig of 840 bp in length
* 54259 54368: gap of 100 bp
* 54369 55223: contig of 861 bp in length
* 55224 55323: gap of 100 bp
* 55324 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 853 bp in length
* 60984 61083: gap of 100 bp
* 61084 61933: contig of 852 bp in length
* 61934 62033: gap of 100 bp
* 62034 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match          100.0%; Score 139; DB 2; Length 59570;
Best Local Similarity 100.0%; Pred. No. 7.6e-32;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCCGGCTCCAGTCTTCTCTGAGTGACTGAACATCAATACAGAGGCGCGGAA 60
|||||
Db 41406 GGGCCCCGGCTCCAGTCTTCTCTGAGTGACTGAACATCAATACAGAGGCGCGGAA 41465

QY 61 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
|||||
Db 41466 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 41525

QY 121 GCCGATCTATAAAGGAA 139
|||||
Db 41526 GCCGATCTATAAAGGAA 41544

RESULT 5
AL359182/c
LOCUS          AL359182          96717 bp    DNA    linear    PRI 11-JAN-2002
DEFINITION    Human DNA sequence from clone RP11-217B7 on chromosome 9, complete
sequence.
ACCESSION    AL359182
VERSION      AL359182.20  GI:18151453
KEYWORDS     HTG.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 96717)
AUTHORS      Skuce,C.
TITLE        Direct Submission
JOURNAL      Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,

```

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jan 15, 2002 this sequence version replaced g1:18121468. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; information on the WormPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone cortigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HP/Chr9> RP11-217B7 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-217B7 it may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RP11-217B7 is at 96717 in this sequence. The true left end of clone RP11-122F10 is at 72980 in this sequence. The true right end of clone RP11-31J20 is at 2000 in this sequence.

FEATURES

```

source      Location/Qualifiers
1..96717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-217B7"
/clone_lib="RP11-11.1"
3238..3278
misc_feature
/note="Sequence from AF275948 sequenced by National Institutes of Health, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA."
84249..84273
misc_feature
/note="Sequence from overlapping clone RP11-122F10 (AC026643). Assembly confirmed by restriction digest."
92050..92163
misc_feature
/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
92411..92557
misc_feature
/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
BASE COUNT  27673 a 21138 c 20380 g 27526 t
ORIGIN
Query Match          100.0%; Score 139; DB 9; Length 96717;
Best Local Similarity 100.0%; Pred. No. 7.3e-32;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCCCGGCTCCAGTCTTCTCTGAGTGACTGAACATCAATACAGAGGCGCGGAA 60
|||||
Db 17416 GGGCCCCGGCTCCAGTCTTCTCTGAGTGACTGAACATCAATACAGAGGCGCGGAA 17357
QY 61 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
|||||
Db 17356 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 17297

```



```

Oy 121 GCCGAATCTATTAAGGAA 139
|||||T|||||T|||||
Db 17296 GCCGAATCTATTAAGGAA 17278

RESULT 6
AC012230
LOCUS
DEFINITION
AC012230 175064 bp DNA linear HTG 22-APR-2000
Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
pieces.
AC012230
VERSION
AC012230.3 GI:7637254
KEYWORDS
HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1M10
Unpublished
2 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckerly,R., Bozulavkiy,L., Bouhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArlano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funk,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2510
Center clone name: L_M_10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145743 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-contigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1003: contig of 1003 bp in length
* 1004 1103: gap of 100 bp
*
1104 2634: contig of 1531 bp in length
2635 2734: gap of 100 bp
2735 4415: contig of 1681 bp in length
4416 4515: gap of 100 bp
4516 5785: contig of 1270 bp in length
5786 5885: gap of 100 bp
5886 7879: contig of 1994 bp in length
7880 7979: gap of 100 bp
7980 9686: contig of 1707 bp in length
9687 9786: gap of 100 bp
9787 12253: contig of 2467 bp in length
12254 12353: gap of 100 bp
12354 15228: contig of 2875 bp in length
15229 15328: gap of 100 bp
15329 17200: contig of 1872 bp in length
17201 17300: gap of 100 bp
17301 20131: contig of 2831 bp in length
20132 20331: gap of 100 bp
20332 22587: contig of 2356 bp in length
22588 22687: gap of 100 bp
22688 25707: contig of 3026 bp in length
25708 25807: gap of 100 bp
25808 28184: contig of 2377 bp in length
28185 28284: gap of 100 bp
28285 31338: contig of 3054 bp in length
31339 31438: gap of 100 bp
31439 34299: contig of 2861 bp in length
34300 34399: gap of 100 bp
34400 38318: contig of 3919 bp in length
38319 38418: gap of 100 bp
38419 42835: contig of 4417 bp in length
42836 42935: gap of 100 bp
42936 45448: contig of 2513 bp in length
45449 45548: gap of 100 bp
45549 48116: contig of 2568 bp in length
48117 48216: gap of 100 bp
48217 52618: contig of 4402 bp in length
52619 52718: gap of 100 bp
52719 56592: contig of 3874 bp in length
56593 56692: gap of 100 bp
56693 59635: contig of 2943 bp in length
59636 59735: gap of 100 bp
59736 63661: contig of 3926 bp in length
63662 63761: gap of 100 bp
63762 68437: contig of 4676 bp in length
68438 68537: gap of 100 bp
68538 71458: contig of 2921 bp in length
71459 71558: gap of 100 bp
71559 76888: contig of 5330 bp in length
76889 76988: gap of 100 bp
76989 82113: contig of 5125 bp in length
82114 82213: gap of 100 bp
82214 88220: contig of 6007 bp in length
88221 88320: gap of 100 bp
88321 93499: contig of 5179 bp in length
93500 93599: gap of 100 bp
93600 97901: contig of 4302 bp in length
97902 98001: gap of 100 bp
98002 103016: contig of 5015 bp in length
103017 103116: gap of 100 bp
103117 109178: contig of 6062 bp in length
109179 109278: gap of 100 bp
109279 117307: contig of 8029 bp in length
117308 117407: gap of 100 bp
117408 124079: contig of 6572 bp in length
124080 124179: gap of 100 bp
124180 131281: contig of 7102 bp in length
131282 131381: gap of 100 bp
131382 138059: contig of 6678 bp in length
138060 138159: gap of 100 bp
138160 145491: contig of 7332 bp in length
145492 145591: gap of 100 bp
145592 157391: contig of 11800 bp in length

```

* 157392 157431: gap of 100 bp
 * 157492 175064: contig of 17573 bp in length.

FEATURES

Source

Location/Qualifiers
 1..175064

/organism="Homo sapiens"

/gb_xref="taxon:9606"

/clone="RP11-1M10"

/clone_lib="3PCI-11 Human Male BAC"

misc_feature

1..1003

/note="assembly_fragment"

misc_feature

1104..2634

/note="assembly_fragment"

misc_feature

2735..4415

/note="assembly_fragment"

misc_feature

4516..5785

/note="assembly_fragment"

misc_feature

5886..7879

/note="assembly_fragment"

misc_feature

7980..9686

/note="assembly_fragment"

misc_feature

9787..12253

/note="assembly_fragment"

misc_feature

12354..15228

/note="assembly_fragment"

misc_feature

15329..17200

/note="assembly_fragment"

misc_feature

17301..20131

/note="assembly_fragment"

misc_feature

20232..22587

/note="assembly_fragment"

misc_feature

22688..25707

/note="assembly_fragment"

misc_feature

25808..28184

/note="assembly_fragment"

misc_feature

28285..31338

/note="assembly_fragment"

misc_feature

31439..34299

/note="assembly_fragment"

misc_feature

34400..38318

/note="assembly_fragment"

misc_feature

38419..42835

/note="assembly_fragment"

misc_feature

42936..45448

/note="assembly_fragment"

misc_feature

45543..48116

/note="assembly_fragment"

misc_feature

48217..52618

/note="assembly_fragment"

misc_feature

52719..56592

/note="assembly_fragment"

misc_feature

56693..59635

/note="assembly_fragment"

misc_feature

59736..63661

/note="assembly_fragment"

misc_feature

63762..68437

/note="assembly_fragment"

misc_feature

68538..71458

/note="assembly_fragment"

misc_feature

clone_end:SP6

vector_side:left

71559..76888

Query Match 100.0%; Score 139; DB 2; Length 175064;

Best Local Similarity 100.0%; Pred. No. 6.8e-32;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCCGGCTCCACGTGCTTTCTGCTGAGTGAACACTACATAACAGAGCGCGGAA 60

|||||

Db 3503 GGGCCCCGGCTCCACGTGCTTTCTGCTGAGTGAACACTACATAACAGAGCGCGGAA 3562

|||||

QY 61 CGGGCGGGGAGGAGGAGACAGGCTTTGACCATAGTACCTCTCGGCTCGGTGCA 120

|||||

Db 3563 CGGGCGGGGAGGAGGAGACAGGCTTTGACCATAGTACCTCTCGGCTCGGTGCA 3622

QY 121 GCCGAATCTATAAAGGAA 139

|||||

Db 3623 GCCGAATCTATAAAGGAA 3641

RESULT 7

LOCUS

AX092589

Sequence 1 from Patent WO0115676.

ACCESSION

AX092589

VERSION

AX092589.1

GI:13444647

KEYWORDS

Human.

SOURCE

Human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 183999)

AUTHORS

Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.

Compositions and methods for modulating hdl cholesterol and

triglyceride levels

JOURNAL

Patent: WO 0115676-A 1 08-MAR-2001;

University of British Columbia (CA) ; Xenon Genetics Inc. (CA)

FEATURES

source

..183999

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT

49549 a 37944 c 41170 g 54950 t

ORIGIN

Query Match 100.0%; Score 139; DB 6; Length 183999;

Best Local Similarity 100.0%; Pred. No. 6.8e-32;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCCGGCTCCACGTGCTTTCTGCTGAGTGAACACTACATAAAGAGCGCGGAA 60

|||||

Db 28631 GGGCCCCGGCTCCACGTGCTTTCTGCTGAGTGAACACTACATAAAGAGCGCGGAA 28690

|||||

QY 61 CGGGCGGGGAGGAGGAGACAGGCTTTGACCATAGTACCTCTCGGCTCGGTGCA 120

|||||

Db 28691 CGGGCGGGGAGGAGGAGACAGGCTTTGACCATAGTACCTCTCGGCTCGGTGCA 28750

|||||

QY 121 GCCGAATCTATAAAGGAA 139

|||||

Db 28751 GCCGAATCTATAAAGGAA 28769

|||||

RESULT 8

AF258623S1

LOCUS

AF258623S1

Sequence 1 from Patent WO0115676.

ACCESSION

AF258623

VERSION

AF258623.2

GI:8677405

KEYWORDS

1 of 4

SEGMENT

Human.

SOURCE

Human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1167)

AUTHORS

Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,

Aouizerat,B.E., Fielding,C.J. and Kane,J.P.

Analysis of hABCL gene 5' end: additional peptide sequence,

promoter region, and four polymorphisms

JOURNAL

Biochem. Biophys. Res. Commun. 271 (2000) in press

REFERENCE

2 (bases 224 to 1167)

AUTHORS

Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,

Aouizerat,B.E., Fielding,C.J. and Kane,J.P.

Direct Submission

JOURNAL

Submitted (20-APR-2000) Cardiovascular Research Institute,

University of California, San Francisco, 505 Parnassus Avenue, San

Francisco, CA 94143-0130, USA

Sequence 1 from Patent WO0115675.					
DEFINITION					
ACCESSION	AX092589				
VERSION	AX092589.1 GI:13444647				
KEYWORDS	human.				
SOURCE					
ORGANISM	<i>Homo sapiens</i>				
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.				
TITLE	Compositions and methods for modulating hdl cholesterol and triglyceride levels				
JOURNAL	Patent: WO 0115676-A 1 08-MAR-2001;				
FEATURES	University of British Columbia (CA); Xenon Genetics Inc. (CA)				
source	1..183999 Location/Qualifiers /organism="Homo sapiens" /db_xref="taxon:9606"				
BASE COUNT	49549 a 37944 c 41170 g 54950 t 386 others				
ORIGIN					
	Query Match	90.6%;	Score 1387.6;	DB 6;	Length 183999;
	Best Local Similarity	97.7%;	Pred. No. 0;		
	Matches 1503;	Conservative	2;	Mismatches	10;
				Indels	23;
				Gaps	
QY	16	GCCTCCACAATGCACCTTCCAGGGCGCTGCTTGG--CCTCTTATSGGTCTGTCTCGTAGTGTT	73		
Dd	27234	GCCTCCACATGCACCTTCAGGGCGCTCTTGGGCCCTCTTCTATGGGTCGTCTCTGAGTGT	27293		
QY	74	CATAGAACCACACTGATGTGAGTAACCTGGGCTTGAG-CGTGCCCTGGAGAPCCCTTGTGACTG	132		
Dd	27294	CATAGAACCACACTGATGTGAGTAACCTGGGCTTGAGCGCTGGAGATCCCTGTGTGACTG	27353		
QY	133	TAGCATGGAGGGGCTTGT-CAGCTGAATCTCTATGACAGGTGGGAGTCTTGGAAAT	191		
Dd	27354	TAGCATGGAGGGGCTTGTCCAGCTGAATCTCTGTAGCAGGTGGGAGTCTTGGAAAT	27413		
QY	192	ATGATGGAGCTGGAGGTGGGAAGAGAAATAGGCTTGGGCGAGACTCTCATGCAACCTCA	251		
Dd	27414	ATGATGGAGCTGGAGGTGGGAAGAGAAATAGGCTTGGGCGAGCTCTCATGCCACCTCA	27473		
QY	252	TTCTGGCCAAAACATCAGGTCAAACCTGTCAAGAGTCTAAATGTGAATCTGCCCTTCAAGT	311		
Dd	27474	TTCTGGCCAAAACATCAGGTCAAACCTGTCAAGAGTCTAAATGTGAATCTGCCCTTCAAGT	27533		
QY	312	GCCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGGGCCCTCCAGTGCACCTTCCAGG	371		
Dd	27534	GCCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGGGCCCTCCAGTGCACCTTCCAGG	27593		
QY	372	GGCTGCTT-GGCCCTCTTACGGGTCTGCTGAGTCTTCTATGAA--TTCGCCCTCAGG	428		
Dd	27594	GGCTGCTTGGGCGCTTCTTACGGGTCTGCTGAGTCTTCTATGAAATCTGTGCTTTCAGG	27653		
QY	429	GCAGATTCTAATTTAGACTCTTCACAGTTTCAACCTGAGTTTGGCCAGATAAGGTGACA	488		
Dd	27654	GCAGATTCTAATTTAGACTCTTCACAGTTTCAACCTGAGTTTGGCCAGATAAGGTGACA	27713		
QY	489	TTTTAGTTTCTGGCTTGATGAATCACTTAAATATTAGACATATGTTGTTCTAGGCTTGCA	548		
Dd	27714	TTTTAGTTTCTGGCTTGATGAATCACTTAAATATTAGACATATGTTGTTCTAGGCTTGCA	27773		
QY	549	TTCTTACTCTTGGCTTTTTTTTTGGCCCTCCAGTCTTTGGGTAGTTTGTCTCCCCCTAC	608		
Dd	27774	TTCTTACTCTTGGCTTTTTTTTTGGCCCTCCAGTCTTTGGGTAGTTTGTCTCCCCCTAC	27832		
QY	609	AGCCAAGCAACACAGATGAAGTTGGAGGTCTGGAGTGGCTACATATTTTACACGACTGC	668		
Dd	27833	AGCCAAGCAACACAGATGAAGTTGGAGGTCTGGAGTGGCTACATATTTTACACGACTGC	27892		
QY	669	AATTCTCTGGCTGCACCTTCCAAATGTTATACAACTAAATACAGTCTGTTGTTTATC	728		
Dd	27893	AATTCTCTGGCTGCACCTTCCAAATGTTATACAACTAAATACAGTCTGTTGTTTATC	27952		

QY	729	ACAGGAGGGCGTGATCAATAA		"GCCTCCCATATTGTCTGTGCT 788
Dd	27953	ACAGGGAGGCGTATCAATAATAAAGAGGGCGTGT-CCAATTGTTCTGTGT 28011		
QY	789	TTTTGTTTGT-TGTTCCTTTT-----TTTTGTTTGTGCGCTCCTCTCTCAATT 839		
Dd	28012	TTTTGTTTGTTCCTTTNNNTNNTNNTNNTGTTTGTGCGCTCCTCTCTCAATT 28071		
QY	840	TATGAAGAAGACAGTAGTAAGATTTCCTTCGGGTCTCTAGAGGACTGGGAGCTCAGG 899		
Dd	28072	TATGAAGAAGACAGTAGTAAGATTTCCTTCGGGTCTCTAGAGGACTGGGAGCTCAGG 28131		
QY	900	CTGGGAATCTCAAAGGCAGTAGTGGCTTATCAAAAATCAAATCCAGTCTGTCTGGGGG 959		
Dd	28132	CTGGGAATCTCAAAGGCAGTAGTGGCTTATCAAAAATCAAATCCAGTCTGTCTGGGGG 28191		
QY	960	AAAAACAAGCAGCCCCATTACCAGAGAGCTGTTCGGCTTCGCCCTCACCCCAGCCCTAGGC 1019		
Dd	28192	AAAAACAAGCAGCCCCATTACCAGAGAGCTGTTCGGCTTCGCCCTCACCCCAGCCCTAGGC 28251		
QY	1020	CCTTGAAAGGAAACAACAAGAACACAAAATGATTGCGCTTCCTGAGGAGAGATTCAGCCTA 1079		
Dd	28252	CCTTGAAAGGAAACAACAAGAACACAAAATGATTGCGCTTCCTGAGGAGAGATTCAGCCTA 28311		
QY	1080	GAGCTCTCTCTCOOCCAAATCCCTCCCTCGGCTGAGGAATCTACAAGAGAAAAAANAAT 1139		
Dd	28312	GAGCTCTCTCTCOOCCAAATCCCTCCCTCGGCTGAGGAATCTACAAGAGAAAAAANAAT 28371		
QY	1140	TCGGGAAGCAGGATTTAGAGGAAGCAAAATCCACTGTGCGCTTGGCTCCCGGGAACGT 1199		
Dd	28372	TCGGGAAGCAGGATTTAGAGGAAGCAAAATCCACTGTGCGCTTGGCTCCCGGGAACGT 28431		
QY	1200	GGACTAGAGAGTCTCGCGCGCAGCCCCCAGCGCTTCCCGCGGCTCTTAGCGCGGC 1259		
Dd	28432	GGACTAGAGAGTCTCGCGCGCAGCCCCCAGCGCTTCCCGCGGCTCTTAGCGCGGC 28491		
QY	1260	GGGCCCCGGCGGGGAAGGGGACGCAGACCGCGGACCTTAGACACCTGCTGTACCTCC 1319		
Dd	28492	GGGCCCCGGCGGGGAAGGGGACGCAGACCGCGGACCTTAGACACCTGCTGTACCTCC 28551		
QY	1320	A----CCCCCACCCACCCACTCCCCCACTCCCTAGATGTGCTGGCGCGGTGAA 1374		
Dd	28552	ANNNCNNCCACCCACCCACTCCCCCACTCCCTAGATGTGCTGGCGCGGTGAA 28611		
QY	1375	CGTCGCGCGTTTAAGGGCGGGCCCCCGGCTCCACGTGTCTTCTGCTGASTGACTGAActA 1434		
Dd	28612	CGTCGCGCGTTTAAGGGCGGGCCCCCGGCTCCACGTGTCTTCTGCTGASTGACTGAActA 28671		
QY	1435	CATAAACAGAGGCCGGAACGGGGCGGGGAGGAGGAGGAGACAGGCTTGTACCGGATAGT 1494		
Dd	28672	CATAAACAGAGGCCGGAACGGGGCGGGGAGGAGGAGGAGACAGGCTTGTACCGGATAGT 28731		
QY	1495	AACCTCTGCCCTCGGTGCAGCCGGAATCTATAAAGGAA 1532		
Dd	28732	AACCTCTCGCTCGGTGCAGCCGGAATCTATAAAGGAA 28769		
<hr/>				
RESULT 5				
AX351031				
LOCUS	AX351031	2893 bp	DNA	linear PAT 06-FEB-2
DEFINITION	Sequence 3 from Patent WO183746.			
ACCESSION	AX351031			
VERSION	AX351031.1			
KEYWORDS	GI:18616387			
SOURCE	human.			
ORGANISM	Homo sapiens			
Eukaryota:	Metazoa;			
Mammalia:	Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE 1	Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P., Brewer,B., Duverger,N., Remaley,A. and Santamarina-Polo,S. Regulatory nucleic acid sequences of the abcl gene Patent: WO 0183746-A 3 08-NOV-2001; JOURNAL			

RESULT 5					
AX351031					
LOCUS	2893 bp	DNA	linear	PAT 06-FEB-2002	
DEFINITION	Sequence 3 from Patent WO183746.				
ACCESSION	AX351031				
VERSION	AX351031.1	GI:18616387			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1				
AUTHORS	Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P., Brewer,B., Duverger,N., Remaley,A. and Santamarina-Pojo,S. Regulatory nucleic acid sequences of the abcl gene Patent: WO 0183746-A 3 08-NOV-2001;				
TITLE					
JOURNAL					

FEATURES
 source
 Location/Qualifiers
 I..2893
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 745 a 681 c 769 g 698 t
 ORIGIN
 Query Match 89.7%; Score 1374.4; DB 6; Length 2893;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1504; Conservative 0; Mismatches 6; Indels 26; Gaps 10;
 QY 16 GCTCCACATGCACTCCAGGCGCTGTTGG--CCTCTTATGGCTCTGCTGAGTGT 73
 DB 1345 GCTCCACATGCACTCCAGGCGCTGTTGGCCCTCTCTATGGCTCTGCTGAGTGT 1404
 QY 74 GATAGAACACATGATGTGAGTACCTGGCTTGG--CGTGGCCCTGGAGATCCGTGAGCT 132
 DB 1405 GATAGAACACATGATGTGAGTACCTGGCTTGGCCCTGGCTGGAGATCCGTGAGCT 1464
 QY 133 TAGCATGGAGGGGCTGTG--CAGCTGAATGCTGTATGACAGTGTGGAGTCTGGAT 191
 DB 1465 TAGCATGGAGGGGCTGTG--CAGCTGAATGCTGTATGACAGTGTGGAGTCTGGAT 1524
 QY 152 ATGATCGAGCTGGAGTGGAGAGAGTATGCTTGGGSCAGCTCTCTCATGCCACCTCA 251
 DB 1525 ATGATCGAGCTGGAGTGGAGAGAGTATGCTTGGGSCAGCTCTCTCATGCCACCTCA 1584
 QY 252 TTTCTGGCCAAACTCAGTCAAACTGTAAGAGTCTAAATGTAATCTGCCCTCAAGT 311
 DB 1585 TTTCTGGCCAAACTCAGTCAAACTGTAAGAGTCTAAATGTAATCTGCCCTCAAGT 1644
 QY 312 GCTTACAAAGTATCTTGTCAAGCTAGAGACCTTGTGGCCCTCCACGTCACCTCCAG 371
 DB 1645 GCTTACAAAGTATCTTGTCAAGCTAGAGACCTTGTGGCCCTCCACGTCACCTCCAG 1704
 QY 372 GCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 430
 DB 1705 GCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1761
 QY 431 AGATTCATATTTAGACTCTTCCAGTCTGACCTGAGTCTTGGCCAGAAATAGGTGACAT 490
 DB 1762 AGATTCATATTTAGACTCTTCCAGTCTGACCTGAGTCTTGGCCAGAAATAGGTGACAT 1821
 QY 491 TAGTTTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 550
 DB 1822 TAGTTTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1879
 QY 551 CCTACTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 610
 DB 1880 CCTACTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1938
 QY 611 CCAAGGGCAACAGATAAGTCTGGAGTCTGGAGTCTGATATTTAGACATATGGTGTAGG 670
 DB 1939 CCAAGGGCAACAGATAAGTCTGGAGTCTGGAGTCTGATATTTAGACATATGGTGTAGG 1998
 QY 671 TTTCTGGCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 730
 DB 1999 TTTCTGGCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2058
 QY 731 AGGAGGCTGATCAATATATGAAATTAAGGGGCTGTGTCACATATGCTGTGTTT 790
 DB 2059 AGGAGGCTGATCAATATATGAAATTAAGGGGCTGTGTCACATATGCTGTGTTT 2117
 QY 791 TTG-----TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 841
 DB 2118 TTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 2177
 QY 842 TGAAGAGAGCAGTAAAGTCTTCTCGGGTCTCTGAGGGACCTGGGAGCTCAGGCT 901
 DB 2178 TGAAGAGAGCAGTAAAGTCTTCTCGGGTCTCTGAGGGACCTGGGAGCTCAGGCT 2237
 QY 902 GGGAACTCCAGGAGTATAGTCTGCTATCAAAAAATCAAAAGTCCAGGTTTGTGGGGGAA 961

DB 2238 GGGATCTCCAGGACGAGTGGCTGCTATCAAAATCAAGTCCAGTCTTGTGGGGGAA 2297
 QY 962 AACAAAAGCAGCCCATATACCCAGAGGAGTCTCCGCTTCCCTCACCCACAGCTTAGGCT 1021
 DB 2298 AACAAAAGCAGCCCATATACCCAGAGGAGTCTCCGCTTCCCTCACCCACAGCTTAGGCT 2357
 QY 1022 TTGAAGGAACAAAGACAAAGACAAATGATGGCTTCTGAGGAGATTCACCTTGA 1081
 DB 2358 TTGAAGGAACAAAGACAAAGACAAATGATGGCTTCTGAGGAGATTCAGCTTGA 2417
 QY 1082 GCTCTCTCTCCCTCAATCCCTCCCTCCGCTGAGGAATCTAACAAAGAAAGAAATTTG 1141
 DB 2418 GCTCTCTCTCCCTCAATCCCTCCCTCCGCTGAGGAATCTAACAAAGAAAGAAATTTG 2477
 QY 1142 CGGAAGCAGAGATTTAGAGGAAGCAATTCACCTGGTCCCTTGGCTCCCGGAGACGTGG 1201
 DB 2478 CGGAAGCAGAGATTTAGAGGAAGCAATTCACCTGGTCCCTTGGCTCCCGGAGACGTGG 2537
 QY 1202 ACTAGAGAGTCTGGGGCGCAGCCCGAGCCGAGCCGCTCCGCGGCTTTAGGCGGCGG 1261
 DB 2538 ACTAGAGAGTCTGGGGCGCAGCCCGAGCCGAGCCGCTTTAGGCGGCGGCTTTAGGCGGCGG 2597
 QY 1262 GCCCGGGCGGGGAAGGGAGCGACAGCCGCGGACCTTAAGACACCTGTGTACCTCCAC 1321
 DB 2598 GCCCGGGCGGGGAAGGGAGCGACAGCCGCGGACCTTAAGACACCTGTGTACCTCCAC 2657
 QY 1322 -----CCCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1376
 DB 2658 CCCCCACCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2717
 QY 1377 TCGCCGCTTTTAAAGGGCGGGCGGCTCCAGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1436
 DB 2718 TCGCCGCTTTTAAAGGGCGGGCGGCTCCAGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2777
 QY 1437 TAAACAGAGCGCGGAACGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
 DB 2778 TAAACAGAGCGCGGAACGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2837
 QY 1497 CCTCTGCGCTCGGTGCGAGCCGAATCTATAAAGGAA 1532
 DB 2838 CCTCTGCGCTCGGTGCGAGCCGAATCTATAAAGGAA 2873
 RESULT 6
 LOCUS AX351029 3231 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 1 from Patent WO01-83746.
 ACCESSION AX351029
 VERSION AX351029.1 GI:18616385
 KEYWORDS
 SOURCE human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 AUTHORS
 Rosier-Montus M.F., Prades C., Lenoire C., Naudin L., Deneffe P.,
 Breuer B., Duverger N., Remaley A. and Santamarina-Fojo S.
 TITLE
 Regulatory nucleic acid sequences of the abcl gene
 JOURNAL
 Patent: WO 0183746-A 1 08-NOV-2001;
 Aventis Pharma S.A. (FR)
 FEATURES
 Location/Qualifiers
 1..3231
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 809 a 773 c 876 g 773 t
 ORIGIN
 Query Match 89.7%; Score 1374.4; DB 6; Length 3231;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1504; Conservative 0; Mismatches 6; Indels 26; Gaps 10;
 QY 16 GCTCCACATGCACTCCAGGCGCTGTTGG--CCTCTTATGGCTCTGCTGAGTGT 73

```

Db 28672 CATAAACAGAGCCGGGACCGGCGGGAGAGGAGGACACAGGCTTTGACCGATAGT 28731
QY 1495 AACCTCTCGCTCGGTGAGCGCGAATCTATAAAGGAACCTAGTCCCGGCAAAAACCCCGT 1554
Db 28732 AACCTCTCGCTCGGTGAGCGCGAATCTATAAAGGAACCTAGTCCCGGCAAAAACCCCGT 28791
QY 1555 AATGCGAGCGAGAGTGTAGTGGCGCGGACCCGAGAGCGGACCGCTTCTCTCC 1614
Db 28792 AATGCGAGCGAGAGTGTAGTGGCGCGGACCCGAGAGCGGACCGCTTCTCTCC 28851
QY 1615 GGGCTGGGCGAGGCGAGGCGGGGAGGCTC 1643
Db 28852 GGGCTGGGCGAGGCGAGGCGGGGAGGCTC 28880

RESULT 5
AX351029
LOCUS
DEFINITION
Sequence 1 from Patent WO183746.
ACCESSION
AX351029
VERSION
AX351029.1
KEYWORDS
GI:18616385
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS
Rosier-Montus,K.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P.,
Brewer,B., Duverger,N., Remaley,A. and Santamarina-Pojo,S.,
Regulatory nucleic acid sequences of the abcl gene
Patent: WO 0183746-A 1 08-NOV-2001;
Aventis pharma S.A. (FR)
FEATURES
Location/Qualifiers
1..3231
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT
805 a 773 c 876 g 773 t
ORIGIN
Query Match
Best Local Similarity 90.4%; Score 1485.4; DB 6; Length 3231;
Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;
QY 16 GGTCCACATGCACTTCCAGGGCTGCTTGG--CTCTTCTATGGGTGTCCTGAGTGT 73
Db 1345 GCCTCCACATGCACTTCCAGGGCTGCTTGGGCTCTTCTATGGGTGTCCTGAGTGT 1404
QY 74 GATAGAACCACTGATGTGATGATGCTGGGCTTGAG--CGTGGCTTGGAGATCCTGTCACIG 132
Db 1405 GATAGAACCACTGATGTGATGATGCTGGGCTTGAGGCTGGCTGGAGATCCTGTCACIG 1464
QY 133 TAGCATGGAGGGGCTGTG--CAGCTGAATGTCTGTATGATGAGTGTGGAGTTCCTGSAAT 191
Db 1465 TAGCATGGAGGGGCTGTG--CAGCTGAATGTCTGTATGATGAGTGTGGAGTTCCTGSAAT 1524
QY 192 ATGATGAGCTGGAGTGGGAGAGAGTGTGCTGGGCTTGAGGCTGGGAGTTCCTGSAAT 1524
Db 1525 ATGATGAGCTGGAGTGGGAGAGAGTGTGCTGGGCTTGAGGCTGGGAGTTCCTGSAAT 1524
QY 252 TTCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAGGT 311
Db 1585 TTCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAGGT 1644
QY 312 GGTACACAGGTATCTTCTCAAGTAGGAGACCTTGTGGCTTCCAGTGCACCTCCAGG 371
Db 1645 GGTACACAGGTATCTTCTCAAGTAGGAGACCTTGTGGCTTCCAGTGCACCTCCAGG 1704
QY 372 GGTCTGCTTCTTACGGTCTCTTCTGAGTCTTCTATGATATCCCTTCAGGCT 430
Db 1705 GGTCTGCTTCTTACGGTCTCTTCTGAGTCTTCTATGATATCCCTTCAGGCT 1761
QY 431 AGATTCATATTTAGACTCTTACAGTCTTACAGTCTTACAGTCTTACAGTCTTACAGT 490

```

```

Db 1762 AGATTCATATTTAGACTCTTTCACAGTTTGACCTGAGTTTGGCCAGAAATAGGTCACATT 1821
QY 491 TAGTTTGTGCTTGTGATGAATGACCTTAATATTTAGACATATGTTGTGTAGGCTGCAATT 550
Db 1822 TAGTTTGTGCTTGTGATGAATGACCTTAATATTTAGAC--ATGGTGTGTAGGCTGCAATT 1879
QY 551 CCTACICTGTGCTTGTGATGAATGACCTTAATATTTAGGCTGTTTGGTGTAGTTTGTCTCCCTACAG 610
Db 1880 CCTACTCTGTGCTTGTGATGAATGACCTTAATATTTAGGCTGTTTGGTGTAGTTTGTCTCCCTACAG 1938
QY 611 CCAAGGCAACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
Db 1939 CCAAGGCAACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1998
QY 671 TTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
Db 1999 TTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2058
QY 731 AGGAGGCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
Db 2059 AGGAGGCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2117
QY 791 TTG-----TTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
Db 2118 TTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2177
QY 842 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
Db 2178 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2237
QY 902 GGAATCTCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
Db 2238 GGAATCTCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
QY 962 AACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
Db 2298 AACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2357
QY 1022 TTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
Db 2358 TTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2417
QY 1082 GCTCTCTCTCCCGCAATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1141
Db 2418 GCTCTCTCTCCCGCAATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2477
QY 1142 CGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
Db 2478 CGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2537
QY 1202 ACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
Db 2538 ACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2597
QY 1262 GCGCGGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1321
Db 2598 GCGCGGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
QY 1322 -----CCCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1376
Db 2658 CCCCCACCCACCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2717
QY 1377 TCGCCCGGTTTAAAGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 1436
Db 2718 TCGCCCGGTTTAAAGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 2777
QY 1437 TAAACAGAGGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1496
Db 2778 TAAACAGAGGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2837
QY 1497 CCTCTGCGCTCGGTGACGCGCGAATCTATAAAGGAAGTGTAGTCCCGGCAAAACCCCGTAA 1556

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 14:33:17 : Search time 92.7366 Seconds
(without alignments)
7959.555 Million cell updates/sec

Title: US-09-596-141c-3

Perfect score: 1643

Sequence: 1 gaattccttgccggtggtc.....caggcaggcgaggagatc 1643

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224532407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2.6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1485.4	90.4	3231	10	US-09-846-456-1 Sequence 1, Appli
2	1394.4	84.9	2893	10	US-09-846-456-3 Sequence 3, Appli
3	102.2	6.2	2893	10	US-09-846-456-3 Sequence 3, Appli
4	102.2	6.2	3231	10	US-09-846-456-1 Sequence 1, Appli
5	91	5.5	221	10	US-09-846-456-4 Sequence 4, Appli
6	43.4	2.6	32190	9	US-09-850-670-255 Sequence 255, App
7	43.4	2.6	32269	9	US-09-860-670-260 Sequence 260, App
8	43	2.6	366	9	US-09-796-692-8332 Sequence 8332, Ap
9	41.8	2.5	251	10	US-09-764-855-6C Sequence 60, Appl
10	41.6	2.5	1974	10	US-09-834-975-994 Sequence 994, App
11	41.4	2.5	2099	10	US-09-764-870-14 Sequence 14, Appl
12	41	2.5	12718	10	US-09-764-877-3972 Sequence 3972, Ap
13	40.6	2.5	545	10	US-09-878-574-4299 Sequence 4299, Ap
14	40.2	2.4	1372	9	US-10-008-118A-19 Sequence -9, Appl
15	40.2	2.4	1372	10	US-09-443-704-19 Sequence 19, Appl
16	40	2.4	16511	10	US-09-764-869-2064 Sequence 2064, Ap
17	39.6	2.4	2756	10	US-09-925-30-351 Sequence 351, App
18	39.4	2.4	335	10	US-09-834-975-433 Sequence 433, App
19	39	2.4	668	10	US-09-731-872-192 Sequence 192, App

```

C 20 39 2.4 2277 10 US-09-917-800A-1513 Sequence 1513, Ap
C 21 38.6 2.3 1338 10 US-09-800-723-44 Sequence 44, Appl
C 22 38.4 2.3 295 10 US-09-764-846-33 Sequence 33, Appl
C 23 38.4 2.3 304 10 US-09-764-846-109 Sequence 109, App
C 24 38.4 2.3 310 10 US-09-864-864-162 Sequence 162, App
C 25 38.4 2.3 2504 10 US-09-729-674-147 Sequence 147, App
C 26 38.4 2.3 32195 10 US-09-764-870-611 Sequence 611, App
C 27 38.4 2.3 32195 10 US-09-764-870-617 Sequence 617, App
C 28 38.4 2.3 32195 10 US-09-764-869-1605 Sequence 1605, App
C 29 38.2 2.3 2509 10 US-09-925-301-540 Sequence 540, App
C 30 38 2.3 1113 10 US-09-834-975-759 Sequence 759, App
C 31 38 2.3 1484 9 US-09-932-598-292 Sequence 292, App
C 32 38 2.3 1484 9 US-09-989-293A-292 Sequence 292, App
C 33 38 2.3 1484 9 US-09-989-735-292 Sequence 292, App
C 34 38 2.3 1484 9 US-09-990-444-292 Sequence 292, App
C 35 38 2.3 1484 9 US-09-989-730-292 Sequence 292, App
C 36 38 2.3 1484 9 US-09-990-436-292 Sequence 292, App
C 37 38 2.3 1484 9 US-09-991-181-292 Sequence 292, App
C 38 38 2.3 1484 9 US-09-993-687-292 Sequence 292, App
C 39 38 2.3 1484 9 US-09-989-734-292 Sequence 292, App
C 40 38 2.3 1484 9 US-09-997-653-292 Sequence 292, App
C 41 38 2.3 1484 9 US-10-174-590-219 Sequence 219, App
C 42 38 2.3 1484 9 US-10-176-758-219 Sequence 219, App
C 43 38 2.3 1484 9 US-10-175-737-219 Sequence 219, App
C 44 38 2.3 1484 10 US-09-983-722-292 Sequence 292, App
C 45 38 2.3 1484 10 US-09-983-723-292 Sequence 292, App

```

ALIGNMENTS

```

RESULT 1
US-09-846-456-1
; Sequence 1, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Gendrine
; APPLICANT: Naudin, Laurent
; APPLICANT: Benefle, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
; TITLE OF INVENTION: Activity and Therapeutic Uses
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-1

```

```

Query Match 50.4%; Score 1485.4; DB 10; Length 3231;
Best local similarity 98.1%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

```

```

QY 16 GGCTCCACATGCCACTTCCAGGGCTGCTTGG--CCTTCTATGGGCTGCTCTGAGTGT 73
Db 1345 GCTCCACATGCCACTTCCAGGGCTGCTTGGGCTCTCTATGCTGCTGCTGAGTGT 1404
QY 74 GATAGAACCACTGATGTGAGTACCTGGCTTGAG-CGTGGCCTGGAGATCCTGTTGACTG 132
Db 1405 GATAGAACCACTGATGTGAGTACCTGGCTTGAGCCCTGGCCTGGAGATCCTGTTGACTG 1464
QY 133 TAGCATGAGAGGGGCTTTGT-CAGCTGAATGTTGATGCAGGTGGTGGGAGTCTTGGAT 191

```

```
Db 1465 TAGCATGGAGGGGCTGTGCGAGCTGAATGTCTGCATGAGGTGGGAGTCTGCAAT 1524
QY 192 ATGATGGAGCTGGAGTGGGAAGAGTAGGCTTGGGCGAGCTCTCTCATGCACTCA 251
Db 1525 ATGATGGAGCTGGAGTGGGAAGAGTAGGCTTGGGCGAGCTCTCTCATGCACTCA 1584
QY 252 TTCGCGCCAAACTCAGGTCAACTGTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGAAGT 311
Db 1585 TTCGCGCCAAACTCAGGTCAACTGTGAAGTCTGAAGTCTGAAGTCTGAAGTCTGAAGT 1644
QY 312 GGCTACAAAGGTATCTTGTCAAGGTAGAGACCTTGTGCGCTCCAGCTGCACCTCCAGG 371
Db 1645 GGCTACAAAGGTATCTTGTCAAGGTAGAGACCTTGTGCGCTCCAGCTGCACCTCCAGG 1704
QY 372 GCGTGGTT--GGGCTCTTACGGGTCTGCTGAGTCTTCTATGATCTCCCTTACGGG 430
Db 1705 GCGTGGTTGGGCTCTTCTACGGGTCTGCTGAGTCTTCTATGATCTCCCTTACGGG 1761
QY 431 AGATTCAATTTAGACTCTTCACAGTTTGCACCTGAGTTTGGCCAGAAATAAGTGCAT 490
Db 1762 AGATTCAATTTAGACTCTTCACAGTTTGCACCTGAGTTTGGCCAGAAATAAGTGCAT 1821
QY 491 TAGTTTGTGGCTTGATGATGACATAATATTTAGACATATGTTGTAGGCTGCAT 550
Db 1822 TAGT-TGTTGGCTTGATGATGACATAATATTTAGAC--ATGGTGTAGGCTGCAT 1879
QY 551 CTAATCTTGGCTTTT-TTGGCCGCTCCAGTGTATTGGGTATTTGCTCCCTTACAG 610
Db 1880 CTAATCTTGGCTTTT-TTGGCCGCTCCAGTGTATTGGGTATTTGCT-CCCTTACAG 1938
QY 611 CCAAGGCAACAGATAAGTTCGAGTCTGGAGTGGCTACATAATTTACAGACTGCA 670
Db 1939 CCAAGGCAACAGATAAGTTCGAGTCTGGAGTGGCTACATAATTTACAGACTGCA 1998
QY 671 TTCTCTGGCTGCACCTTCAAAATGATACAAACTAAATCAAGTCTGTGTTTATCAC 730
Db 1999 TTCTCTGGCTGCACCTTCAAAATGATACAAACTAAATCAAGTCTGTGTTTATCAC 2058
QY 731 AGGAGGCTGATCAATAATGAATTAATAAGGGGCTGGTCCCATATTTCTGTGTT 790
Db 2059 AGGAGGCTGATCAATAATGAATTAATAAGGGGCTGGT-CCATATTTCTGTGTT 2117
QY 791 TTG-----TTTGTGTTGTTCTTTCTTTGTTTGTGTTGTTGTTGTTGTTGTTGTT 841
Db 2118 TTGTTGTTGTTTGTGTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2177
QY 842 TGAAGAGAGCAGTAAGATGTTTCTCTCGGTCCTCTGAGGACCTGGGAGCTCAGGCT 901
Db 2178 TGAAGAGAGCAGTAAGATGTTTCTCTCGGTCCTCTGAGGACCTGGGAGCTCAGGCT 2237
QY 902 GGAATCTCCAGGCGAGTAGTTCGCTATCAAAATCAAAAGTCCAGTTTGTGGGGGAA 961
Db 2238 GGAATCTCCAGGCGAGTAGTTCGCTATCAAAATCAAAAGTCCAGTTTGTGGGGGAA 2297
QY 962 AACAAAGCAGCCCATACCCAGAGTCTGCGGCTTCCCTTCCCTTCCCTTCCCTTCCCT 1021
Db 2298 AACAAAGCAGCCCATACCCAGAGTCTGCGGCTTCCCTTCCCTTCCCTTCCCTTCCCT 2357
QY 1022 TTCAAGGAAACAAAGACAGCAAAATGATGCGTCTCAGGAGATTCAGGCTAGA 1081
Db 2358 TTCAAGGAAACAAAGACAGCAAAATGATGCGTCTCAGGAGATTCAGGCTAGA 2417
QY 1082 GCTCTCTCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1141
Db 2418 GCTCTCTCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2477
QY 1142 CGGAAGCAGGATTCAGAGGAGCAAAATTCACCTGTTGCTGCTGCTGCTGCTGCTGCTG 1201
Db 2478 CGGAAGCAGGATTCAGAGGAGCAAAATTCACCTGTTGCTGCTGCTGCTGCTGCTGCTG 2537
QY 1202 ACTAGAGATCTGCGGCGAGCCGAGCCAGCTTCCCTGCGGCTTCCCTGCGGCTTCCCT 1261
```

RESULT 2

```
US-09-846-456-3
; Sequence 3, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Catherine
; APPLICANT: Naudin, Laurent
; APPLICANT: Denefle, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: FOJO, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-3
```

Query Match: 84.9%; Score 1394.4; DB 10; Length 2893;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

```
QY 16 GCTCCACATGACATCCAGGCGCTGTGG--CTTCTATGGGCTGCTGCTGAGTGT 73
Db 1345 GCCTCCACATGACATCCAGGCGCTGTGG--CTTCTATGGGCTGCTGCTGAGTGT 1404
QY 74 GATGAACACCTGATGATGAGTACCTGGGCTTGAG--CGTGGGCTGGAGATCTGTTGACTG 132
Db 1405 GATGAACACCTGATGATGAGTACCTGGGCTTGAG--CGTGGGCTGGAGATCTGTTGACTG 1464
QY 133 TAGCATGGAGGGGCTTGT-CAGCTGAATGCTGTATGACAGTGGTGGGAGTCTTGGAA 191
```



```

1465 TAGCAATGGAGGGGGCTGTGTCAGCTGATGTCTCGATCGCAGGTGGTGGAGTCTTGGAAAT 1524
1492 ATGATGGAGCTGGAGGGGGGAGAGAGTAGCTGTGGGGAGCTCTCTCATGCAACCTCA 251
1525 ATGATGGAGCTGGAGGGGGGAGAGAGTAGCTGTGGGGAGCTCTCTCATGCAACCTCA 1584
252 TTCTGGCCAAAACCTCAGCTCAAACTGTGAAGAGCTTAAAGTGTGATCTGAGCTTCAAGGT 311
1585 TTCTGGCCAAAACCTCAGCTCAAACTGTGAAGAGCTTAAAGTGTGATCTGAGCTTCAAGGT 1644
312 GGCTACAAAGGCTATCTTGTCAAGGTAGGAGACCTTGTGGCTCCACGCTGACCTTCCAGG 371
1645 GGCTACAAAGGCTATCTTGTCAAGGTAGGAGACCTTGTGGCTCCACGCTGACCTTCCAGG 1704
372 GGCTGCTT-GGCTCTCTCTACGGGTC-GTCCGTGAGTCTCTATGAATCTCCCTTCAGGGC 430
1705 GGCTGCTTGGGCTCTCTCTACGGGTCGTCTGAGTCTCTATGAATCTCCCTTCAGGGC 1761
431 AGATTCTATATTAGACTCTTCCAGCTTGTACCTGAGCTTTGGCCAGAAATAGGTGACATT 490
1762 AGATTCTATATTAGACTCTTCCAGCTTGTACCTGAGCTTTGGCCAGAAATAGGTGACATT 1821
491 TAGTTTGTGTGGCTTGAGCAATGACTTAAATATTATAGACATAAGGTGTAGGCTTGCATT 550
1822 TAGTTTGTGTGGCTTGAGCAATGACTTAAATATTATAGACATAAGGTGTAGGCTTGCATT 1879
551 CCTACTCTTGGCTTTTGTGGCCCTCCAGTGTGTGGGTAGTGTGTGGTCTTCCCTCCAG 610
1880 CCTACTCTTGGCTTTTGTGGCCCTCCAGTGTGTGGGTAGTGTGTGGTCTTCCCTCCAG 1938
611 CCAAGGCAACAGATAAGTTGGAGTCTGGAGTGGGTACATAATTTACACGACTGCAA 670
1939 CCAAGGCAACAGATAAGTTGGAGTCTGGAGTGGGTACATAATTTACACGACTGCAA 1998
671 TTCTCTGCTGCATCTTCAAAAGTATACAAACTAAATACAAAGTCTTGTGTGTATACAC 730
1999 TTCTCTGCTGCATCTTCAAAAGTATACAAACTAAATACAAAGTCTTGTGTGTATACAC 2058
731 AGGAGGCTGATCAATAATATGAATTAAGAGGGGCTGTGCCATATTTCTCTCTT 790
2059 AGGAGGCTGATCAATAATATGAATTAAGAGGGGCTGTGCCATATTTCTCTCTT 2117
791 TGTG-----TTGTTTGTCTTTTGTGTTTGTGCTCTCTCTCTCTCAATTA 841
2118 TGTGTTCTTTCTTTTGTGTTTGTGTTTGTGTTTGTGCTCTCTCTCTCTCAATTA 2177
842 TGAAGAGAGCAAGTAAGATCTTCTCTCGGTCTCTGAGGAGCTGGGAGCTCAGGCT 901
2178 TGAAGAGAGCAAGTAAGATCTTCTCTCGGTCTCTGAGGAGCTGGGAGCTCAGGCT 2237
902 GGAATCTCCAGGCAAGTAGTGTGCTTATCAAAATCAAAAGTCCAGTTTGTGGGGGAA 961
2238 GGAATCTCCAGGCAAGTAGTGTGCTTATCAAAATCAAAAGTCCAGTTTGTGGGGGAA 2297
962 ACAAAGAGAGCCATTTACCCAGGAGTCTCCGCTTCCCTTCCCTTCCCTTCCCTTCCCT 1021
2298 ACAAAGAGAGCCATTTACCCAGGAGTCTCCGCTTCCCTTCCCTTCCCTTCCCTTCCCT 2357
1022 TGAAGAGAGCAAGTAAGATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1081
2358 TGAAGAGAGCAAGTAAGATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2417
1082 GCTCTCTCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1141
2418 GCTCTCTCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2477
1142 CGAAGAGAGGATTTAGAGAGCAATTTCCAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1201
2478 CGAAGAGAGGATTTAGAGAGCAATTTCCAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2537
1202 ACTAGAGAGTCTGGGGGAGCCCGAGCCAGGCTTCCCGGCTCTTGTAGCGCGGCGG 1261
```

```

2538 ACTAGAGAGTCTGGGGGAGCCAGCCCGAGCCCGAGCCCGGCGTTCCTAGCGCGGCGG 2597
1262 GCGGGGGGGGAGGAGGAGCGACAGCCGCGGACCCCTAAGACACCTGTGTATCCCTCCAC 1321
2598 GCGGGGGGGGAGGAGGAGCGACAGCCCGGAGCCCTAAGACACCTGTGTATCCCTCCAC 2657
1322 -----CCCCACCCACCCACCTCCGCCAACTCCTAGATGTGTGCTGGGCGGTGAAG 1376
2658 CCGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2717
1377 TCGCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1436
2718 TCGCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2777
1437 TAAACAGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
2778 TAAACAGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2837
1497 CCTCTGGCTCGGTGCAGCCCAATCTATATAAGGAAGTATAGTCCCGGCAAAACCCC 1552
2838 CCTCTGGCTCGGTGCAGCCCAATCTATATAAGGAAGTATAGTCCCGGCAAAACCCC 2893

RESULT 3
US-09-846-456-3/C
; Sequence 3, Application US/09845456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrine
; APPLICANT: Naudin, Laurent
; APPLICANT: Benefie, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; PRIOR FILING DATE: 2001-05-02
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-3

Query Match 6.2%; Score 102.2; DB 10; Length 2893;
Best Local Similarity 65.9%; Pred. No. 1.5e-17;
Matches 164; Conservative 0; Mismatches 83; Indels 2; Gaps 1;

QY 241 ATGCCACTCTTCTGCGCAAACTCAGGTCAACTGTGAAGAGTCTAATGTGAATCTG 300
Db 1820 ATGTACCTTAT-CTGCGCAAACTCAGGTCAACTGTGAAGAGTCTAATGTGAATCTG 1761
QY 301 CCCTCAAGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCTCCACGT 360
Db 1760 CCCTGAGGATTCATAGAAGACTCAGGACAGACCCGTAGAGAGGCCCAAGGAGGCCCTG 1701
QY 361 GCATTCGAGGGCTGCTTGGCTCTTCTACGGTCTTCTCTGAGTCTTCTATTAATCTC 420
Db 1700 GAAGTGACCTGGAGGC--CACAAAGTCTCTCTACCTTGCACAAAGATACCTTTTGTGCCAC 1643
QY 421 CCTTCAGGCGAGATTCATATTAGACTCTTCAGATTTGACCTGTGAGTTTGGCCAGAATA 480
Db 1642 CTTGAGGCGAGATTCACATTAGACTCTTCAGATTTGACCTGTGAGTTTGGCCAGAATG 1583
QY 481 AGGTGACAT 489
```

```
Db 1582 AGGTGGCAT 1574
|||||
RESULT 4
US-09-846-456-1/c
; Sequence 1, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrine
; APPLICANT: Naudin, Laurent
; APPLICANT: Desefle, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying It
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-1

Query Match 6.2%; Score 102.2; DB 10; Length 3231;
Best Local Similarity 65.9%; Pred. No. 1.6e-17;
Matches 164; Conservative 0; Mismatches 83; Indels 2; Gaps 1;

Qy 241 ATGCCACCTCATTTGCGCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 300
|||||
Db 1820 ATGTCACCTTATTCGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 1761
|||||
Qy 301 CCTCTCAAGTGGGTACAAAGTATCTTTGTCAAGGTAGGAGACCTTGTGGCCCTCCAGT 360
|||||
Db 1760 CCTTGAGGATTCATAGAGAGTCTCAGGACAGCCGCTAGAGAGCCCAAGCCCTCTG 1701
|||||
Qy 361 GCATTCACGGCCCTGTGTCCTCTTCTAGGGTCTCTCTGAGTCTCTCTATGAATCTC 420
|||||
Db 1700 GAAGTGCACGTGGAGGC--CACAGGTCTCTACCTTACACAAAGTACCTTTGTAGCCAC 1643
|||||
Qy 421 CCTTCAGGCGAGATTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATA 480
|||||
Db 1642 CTTGAGGGCGAGATTCACATTAGACTCTTCACAGTTTGAOCTGAGTTTGGCCAGATG 1583
|||||
Qy 481 AGGTGACAT 489
|||||
Db 1582 AGGTGGCAT 1574

RESULT 5
US-09-846-456-4
; Sequence 4, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrine
; APPLICANT: Naudin, Laurent
; APPLICANT: Desefle, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying It
```

```
; TITLE OF INVENTION: Activity and Therapeutic Uses
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-4

Query Match 5.5%; Score 91; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 3.7e-15;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1553 GTAATTCGAGCGAGAGTGAATGGGCGGAGCCGCGAGAGCGGAGCCGAGCCCTTCTCTC 1612
|||||
Db 1 GTAATTCGAGCGAGAGTGAATGGGCGGAGCCGCGAGAGCGGAGCCGAGCCCTTCTCTC 60
|||||
Qy 1613 CCGGCTCGGCGAGGCGAGGCGGCGGAGCTC 1643
|||||
Db 61 CCGGCTCGGCGAGGCGGCGGAGGAGCTC 91
|||||

RESULT 6
US-09-860-670-255
; Sequence 255, Application US/09850670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 32190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-255

Query Match 2.6%; Score 43.4; DB 9; Length 32190;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 62; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 742 TCAATATANTCAATTAAGAGGGGCTGCTCCCATATTGTCTGTCTGTGTTTGTGTTG 801
|||||
Db 25616 TAAAGAAACACAGAGAGAGAGGCTGATCCCAAGCTACAGGTTTTTTTGTGTTG 25675
|||||
Qy 802 TTTCTTTTGTGTTTGTGCTCCCTCCCTCCCT 834
|||||
Db 25676 TTTGTTTGTGTTTGTGAGACACCTCGCTCT 25708
|||||

RESULT 7
US-09-860-670-260
; Sequence 260, Application US/09850670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
```

```
; LENGTH: 32249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-260

Query Match      2.6%; Score 43.4; DB 9; Length 32249;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 62; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 742 TCATATATGAATTAAGAGGGGCTGCTCCCATATGTTCTGTGTTTGTGTTGTTG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18834 TAAAGAAAACAGACAGAGAGAGGCTGATCCCAAGCTACAGGGTTTCTTTGTTG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 802 TTCTTTTGTGTTTGTGCGCTCCCTCTCT 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 TTGTTGTTGTTTGGAGACAGTCTCGCTCT 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-09-796-692-8332/c
; Sequence 8332, Application US/09796692
; Publication No. US20020198352A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2003-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/228,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8332
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (252)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (255)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8332

Query Match      2.6%; Score 43; DB 9; Length 366;
Best Local Similarity 65.6%; Pred. No. 0.045;
Matches 61; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
Qy 742 TCATATATGAATTAAGAGGGGCTGCTCCCATATGTTCTGTGTTTGTGTTGTTG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 TAAAGAAAACAGACAGAGAGGCTGATCCCAAGCTACAGGGTTTCTTTGTTGTTG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 802 TTCTTTTGTGTTTGTGCGCTCCCTCTCT 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 TTGTTGTTGTTTGGAGACAGTCTCGCTCT 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-764-855-60/c
; Sequence 60, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4110
; CURRENT APPLICATION NUMBER: US/09764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-855-60

Query Match      2.5%; Score 41.8; DB 10; Length 251;
Best Local Similarity 65.6%; Pred. No. 0.083;
Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 772 CCATATGTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 CCCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 832 TCTCAATTTATGAAGAGAGCAGTAAAGATGTC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 TTTCAATTTTCAATAATAAATTTAAATCTGTAC 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-834-975-994
; Sequence 994, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Jeffrey
; APPLICANT: Lillie, James
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1045
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 994
; LENGTH: 1974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-834-975-994

Query Match      2.5%; Score 41.6; DB 10; Length 1974;
Best Local Similarity 62.5%; Pred. No. 0.31;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
Qy 717 TGTGTTTATCACAGGGAGCTGATCAATATGAAATTAAAGGGGCTGGTCCCAT 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,471A
; FILING DATE: 04-APR-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 27373/32742A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US91-06532-1
; US-09-283-471A-36

```

```

Query Match          2.4%; Score 38.8; DB 4; Length 1327;
Best Local Similarity 49.2%; Pred. No. 0.33;
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 1171 CCACGTGGCCCTTGGCTGCGGGGAACTGGACTAGAGAGTCTGGGCGGCGAGCCCGGAGC 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 CGCAGAGACACCTGGCGGCTTGGCTGCGGGGAACTGGACTAGAGAGTCTGGGCGGCGAGC 1230

QY 1231 CCAGCGCTTCCCGCGCGTCTTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 667

QY 1291 CGGACCCCTAAGACACCTGCTGTACCTCCACCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CGACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 727

QY 1348 TCCCTAGATGTCTGGTGGCGGCTGAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 TCCGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 787

QY 1408 CGTGTCTTCTGCTGAGTGACTGAAC 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813

```

```

RESULT 7
PCT-US91-06532-1
; Sequence 1, Application PC/TJUS9106532
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
; TITLE OF INVENTION: Vaccines and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

```

```

; ADDRESSEE: Bicknell
; STREET: Two First National Plaza Suite 2100
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06532
; FILING DATE: 19910910
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27373/8235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1335 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US91-06532-1

```

```

Query Match          2.4%; Score 38.8; DB 5; Length 1335;
Best Local Similarity 49.2%; Pred. No. 0.33;
Matches 131; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 1171 CCACGTGGTCCCTTGGCTGCGGGGAACTGGACTAGAGAGTCTGGGCGGCGAGCCCGGAGC 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 CGCAGAGACACCTGGCGGCTTGGCTGCGGGGAACTGGACTAGAGAGTCTGGGCGGCGAGC 607

QY 1231 CCAGCGCTTCCCGCGCGTCTTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 667

QY 1291 CGGACCCCTAAGACACCTGCTGTACCTCCACCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CGACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 727

QY 1348 TCCCTAGATGTCTGGTGGCGGCTGAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 TCCGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 787

QY 1408 CGTGTCTTCTGCTGAGTGACTGAAC 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813

```

```

RESULT 8
US-09-149-476-208/c
; Sequence 208, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/340,333
; EARLIER FILING DATE: 1997-03-07

```

[illegible]


```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163..4099
; US-09-041-886-18

Query Match 2.3%; Score 38.6; DB 4; Length 4481;
Best Local Similarity 52.1%; Pred. No. 0.72;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 1187 CTGCGGGAACGTGACTAGAGAGTCTGGGGGCGCAGCCCCGAGCCAGCGCTTCGCGGCG 1246
DB 194 CGCGGTGCGCCACCGAGTCTCCCGCTTCGCGGCGCAGCGGTGGCGCGGGTGGCGCTCGC 253
QY 1247 GCTTAGCGCGCGCGCGCGCGGGAAGGGGAGCGAGACCGCGGAGCCCTAAGACACC 1306
DB 254 TCACGGCGCGCGCGGCGGAGCGGCGGGGGTGGCGGGGGTGGCGGGCGCGGACCGTATC 313
QY 1307 TCGTGTACCTCCACCCCGACCCCGCCACCTCCCGCCCACTCCCGCCCACTCC 1351
DB 314 COTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 358

RESULT 11
US-08-709-838-1
; Sequence 1, Application US/08/709838
; Patent No. 5140054
; GENERAL INFORMATION:
; APPLICANT: Loetscher, Marcel
; APPLICANT: Moser, Bernhard
; TITLE OF INVENTION: IP-10/MIG RECEPTOR DESIGNATED CXCR3,
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 31-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/709,838
; APPLICATION NUMBER: US 08/709,838
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook Esq., David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: TK-96-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1670 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..1172
; US-08-829-839-1

;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1670 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..1172
; US-08-709-838-1

Query Match 2.3%; Score 38.2; DB 3; Length 1670;
Best Local Similarity 59.8%; Pred. No. 0.55;
Matches 54; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 714 TCCGTGTTTATACACAGGAGCGTGCATATATATGAAATTAAGATTAAGAGGGGCGTGGTCC 773
DB 1562 TCTTTTATTTTATGCTCTAAATCTCGCTTAAAGCTTTCAATAAACAAGATCGTCAGGA 1621
QY 774 CAATATGCTCTGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 820
DB 1622 CCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1668

RESULT 12
US-08-829-839-1
; Sequence 1, Application US/08829839
; Patent No. 6184358
; GENERAL INFORMATION:
; APPLICANT: Loetscher, Marcel
; APPLICANT: Moser, Bernhard
; APPLICANT: Qin, Shixin
; APPLICANT: Mackay, Charles R.
; TITLE OF INVENTION: IP-10/MIG RECEPTOR DESIGNATED CXCR3,
; TITLE OF INVENTION: ANTIBODIES, NUCLEIC ACIDS, AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 31-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/709,838
; APPLICATION NUMBER: US 08/709,838
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook Esq., David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: TK-96-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1670 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..1172
; US-08-829-839-1
```



```
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
```

Query Match 2.3%; Score 37.8; DB 4; Length 162450;
Best Local Similarity 52.2%; Pred. No. 8.4;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
QY 586 TCACAAATGATACAAACTAAATACAGCTCCTGTCGTTTTTATCAGAGGAGGCTGATCAA 745
DB 109593 TAAACATTTTITACTATATATCAATAAGCCCCATTTTATTTCCCTCCCTGTTGCT 109534
QY 746 TATAATGAATTAAGGGGCTGTCATATATGTCGTTGTTGTTGTTGTTGTTGTTGTTGTT 805
DB 109533 CTGATATCTGTGCCAAAGTCACAGTACGCTAAATGATTTTITTTTTTTTTTTT 109474
QY 806 TTTTITGTTTGTGGCTCCTTCCTCTCAATTTAAG 846
DB 109473 TTTTITGTTTGTGGCTCCTTCCTCTCAATTTAAG 846
RESULT 14
US-08-018-977C-4/c
; Sequence 4, Application US/08018577C
; Patent No. 5686601
; GENERAL INFORMATION:
; APPLICANT: Weber, Peter C.
; TITLE OF INVENTION: DNA Sequences Encoding Mutant Antiviral
; TITLE OF INVENTION: Regulatory Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Pennsylvania State University
; STREET: 113 Technology Park
; CITY: University Park
; STATE: PA
; COUNTRY: US
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS 6.22
; SOFTWARE: Microsoft Word for Windows 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/018.977C
; FILING DATE: 18-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/726,071
; FILING DATE: 05-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas J. Monahan
; REGISTRATION NUMBER: 29,835
; REFERENCE/DOCKET NUMBER: 91-1039/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (814) 865-6277
; TELEFAX: (814) 865-3591
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 bases
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-018-977C-4
```

Query Match 2.3%; Score 37.6; DB 1; Length 936;
Best Local Similarity 54.3%; Pred. No. 0.59;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1212 CTGCGCGCAGCCGCCAGCCAGCGCTCCCGCGGCTTTAGCGCGGGGGCGGGCGG 1271
DB 785 CTGCGAGCAGCCGCCCTCTCTCGACAGAGCGGGGCGGGCGGGCGGCGG 726
QY 1272 GCGAAGGGAGCCAGACCCCGGAGCCCTTAAGACACCTGCTGTACCTCCACCCACCCCA 1331
DB 725 GAGGGGGCGCGCGCGGGAGGCGCGTCCACCCCTCCACCGCGGGCCCCCGGAGCGG 666
QY 1332 CCCACCTCCCGCCCACTCCC 1351
DB 665 CGCGCCACCGTCGCAGCGCG 646

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:46:48 : Search time 418.41 Seconds
(without alignments)
8245.850 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1_1532

Perfect score: 1532
Sequence: 1 gaattccctgtgtggctc.....accggaatctataaaaggaa 1532

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1532	100.0	1643	22 AAF24703	Nucleotide sequenc
2	1532	100.0	1643	22 AAF24703	Nucleotide sequenc
3	1387.6	90.6	183999	22 AAF92831	Human ABC1 genomic
4	1374.4	89.7	2910	24 AAD37267	Human ABC1 transcr
5	1374.4	89.7	3231	24 AAD37265	Human ABC1 transcr
6	1035.4	67.6	1197	24 ABL58400	Human large ATP-bi
7	102.2	6.7	2910	24 AAD37267	Human ABC1 transcr
8	102.2	6.7	3231	24 AAD37265	Human ABC1 transcr
9	97	6.3	1643	22 AAF24681	Nucleotide sequenc

C 10	97	6.3	1643	22 AAF24703	Nucleotide sequenc
C 11	92.6	6.0	183999	22 AAF24703	Human ABC1 genomic
C 12	51.4	3.4	474	24 AAF73478	Bovine embryonic g
C 13	45.6	3.0	6116	24 AAF80129	Human chemically m
C 14	44.2	2.9	456	22 AAF80464	Human polynucleoti
C 15	43.8	2.9	385	22 AAF83199	Human polynucleoti
C 16	43.8	2.9	400	22 AAF83631	Human polynucleoti
C 17	43.4	2.8	401	22 AAF85424	Human polynucleoti
C 18	43.4	2.8	772	22 AAF803293	Human cDNA clone (
C 19	43.4	2.8	2339	22 AAF16964	Human cDNA sequenc
C 20	43.4	2.8	10119	22 AAF87559	Human immune/haema
C 21	43.4	2.8	14983	22 AAF73075	Human immune/haema
C 22	43.4	2.8	14983	22 AAF87542	Human immune/haema
C 23	43.4	2.8	32190	22 AAF162927	Human genomic DNA
C 24	43.4	2.8	32249	22 AAF162932	Human immune/haema
C 25	43.4	2.8	39068	22 AAF71820	Human immune/haema
C 26	43.4	2.8	39068	22 AAF73078	Human immune/haema
C 27	43.4	2.8	39068	22 AAF85294	Human immune/haema
C 28	43.4	2.8	39068	22 AAF87544	Human immune/haema
C 29	43.4	2.8	39110	22 AAF71925	Human immune/haema
C 30	43.4	2.8	39110	22 AAF73087	Human immune/haema
C 31	43.4	2.8	39110	22 AAF87555	Human immune/haema
C 32	43.4	2.8	45300	22 AAF73079	Human immune/haema
C 33	43.4	2.8	45300	22 AAF87547	Human immune/haema
C 34	43.2	2.8	421	22 AAF89084	Human polynucleoti
C 35	43.2	2.8	466	22 AAF182085	Human polynucleoti
C 36	43.2	2.8	595	23 ABV16167	Human prostate exp
C 37	43	2.8	470	22 AAF808333	Human cDNA clone (
C 38	42.6	2.8	620	21 AAC59594	Human secreted pro
C 39	42.4	2.8	386	22 AAF92146	Human polynucleoti
C 40	42.4	2.8	389	22 AAF185397	Human polynucleoti
C 41	42.4	2.8	5690	22 AAF81236	Human immune/haema
C 42	42.4	2.8	5690	22 AAF81239	Human immune/haema
C 43	42.2	2.8	429	22 AAF87584	Human polynucleoti
C 44	42.2	2.8	24259	22 AAF66691	Tumour suppressor
C 45	42	2.7	387	22 AAF69088	Human polynucleoti

ALIGNMENTS

RESULT 1

AAF24681
ID AAF24681 standard; DNA; 1643 BP.

XX AAF24681;

XX 20-APR-2001 (first entry)

DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

KW Human; adenosine triphosphate binding cassette protein 1; ABC1;

KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;

KW chromosome 9q22-q31; heart disease; hypercholesterolemia;

KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078972-A2.

XX 28-DEC-2000.

PD 16-JUN-2000; 2000WC-US16765.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-13782/14.

DR

KW apolipoprotein-mediated mobilisation; cholesterol; tanger disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078971-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000MO-US15591.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVH-) CV THERAPEUTICS INC.

XX (UNIW) UNIV WASHINGTON.

XX Lawn RM, Waide D, Oram JF, Garvin M;

XX WPI; 2001-137811/14.

XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
 PT polynucleotides and polypeptides, useful for treatment of heart disease
 PI and other disorders associated with hypercholesterolemia and
 PT atherosclerosis.

XX Disclosure; Page 138-139; 21pp; English.

XX The present sequence represents the 5' flanking region of the human
 CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
 CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
 CC variety of substrates across the plasma membrane. ABC1 is a pivotal
 CC protein in the apolipoprotein-mediated mobilisation of intracellular
 CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
 CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
 CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
 CC are useful for developing pharmaceutical agents for the treatment of
 CC heart disease and other disorders associated with hypercholesterolemia
 CC and atherosclerosis. The genes are useful for developing screening assays
 CC to screen for compounds that regulate the expression of genes associated
 CC with cholesterol transport. The genes and proteins are also useful for
 CC are also useful as diagnostic indicators of cardiovascular disease and
 CC other disorders associated with hypercholesterolemia.

XX Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 100.0%; Score 1532; DB 22; Length 1643;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTGGTCCACATGCACTCCAGGGCGCTGCTGGCTCTCTATGGGTC 60

DB 1 GAATTCCTGCTGGTCCACATGCACTCCAGGGCGCTGCTGGCTCTCTATGGGTC 60

QY 61 TGTCTGAGTGTGATAGAACCACTGATGTCAGTACCTGGGCTTCAGGGCGGCGTGGAGA 120

DB 61 TGTCTGAGTGTGATAGAACCACTGATGTCAGTACCTGGGCTTCAGGGCGGCGTGGAGA 120

QY 121 TCCTGTTGACTGTAGCTAGGAGGGGCTTGCAGTGAATGTCGTATGCAAGTGGTGGG 180

DB 121 TCCTGTTGACTGTAGCTAGGAGGGGCTTGCAGTGAATGTCGTATGCAAGTGGTGGG 180

QY 181 AGTCTGCAATATGATGAGTGGAGTGGGAAGAGATAGGCTTGGGCGAGCTCTCTC 240

DB 181 AGTCTGCAATATGATGAGTGGAGTGGGAAGAGATAGGCTTGGGCGAGCTCTCTC 240

QY 241 ATGCCACCTCATCTTGGCCAAACTCAGGTCAAACTGTSAAGAGTCTAAATGTGAATCTG 300

DB 241 ATGCCACCTCATCTTGGCCAAACTCAGGTCAAACTGTSAAGAGTCTAAATGTGAATCTG 300

QY 301 CCCTTCAGGTGGCTACAAAGTATCTTTGTCAAGTATAGGACACCTTGTGGCCTCCACGT 360

DB 301 CCCTTCAGGTGGCTACAAAGTATCTTTGTCAAGTATAGGACACCTTGTGGCCTCCACGT 360

DB 301 CCCTTCAGGTGGCTACAAAGTATCTTTGTCAAGTATAGGACACCTTGTGGCCTCCACGT 360

QY 361 GCACCTCCAGGGCGCTGCTGGGCTCTCTACGGGTCTGCTCCTGAGTCTTCTATGAATCTC 420

DB 361 GCACCTCCAGGGCGCTGCTGGGCTCTCTACGGGTCTGCTCCTGAGTCTTCTATGAATCTC 420

QY 421 CCTTCAGGGCAGATTCATATTTAGACTCTTCCAGCTTTGACCTGAGTTTGGCCAGAAATA 480

DB 421 CCTTCAGGGCAGATTCATATTTAGACTCTTCCAGCTTTGACCTGAGTTTGGCCAGAAATA 480

QY 481 AGGTGACATTTAGTTGGCTTGGCTGATGATGACTTAAATATTTAGACATATGGTGTGA 540

DB 481 AGGTGACATTTAGTTGGCTTGGCTGATGATGACTTAAATATTTAGACATATGGTGTGA 540

QY 541 GGCTCGCATCTCTAGCTCTTGGCTTTTGGCCCTCCAGTCTTTTGGGTAGTTTGGT 600

DB 541 GGCTCGCATCTCTAGCTCTTGGCTTTTGGCCCTCCAGTCTTTTGGGTAGTTTGGT 600

QY 601 CCCCTACAGCCAAAGSCAAACAGATAAGTTGGAGTCTGAGTGGCTACATAATTTTAC 660

DB 601 CCCCTACAGCCAAAGSCAAACAGATAAGTTGGAGTCTGAGTGGCTACATAATTTTAC 660

QY 661 AGGACTGCAATTTCTGGCTGCACTTCAAAATGTATACAACTAAATCAAGTCCGTG 720

DB 661 AGGACTGCAATTTCTGGCTGCACTTCAAAATGTATACAACTAAATCAAGTCCGTG 720

QY 721 TTTTATPCACAGGAGGCTGATCAATATATGAATTAAGTAAAGGGGGTGGTCCCATATTC 780

DB 721 TTTTATPCACAGGAGGCTGATCAATATATGAATTAAGTAAAGGGGGTGGTCCCATATTC 780

QY 781 TTTCTGTCTTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 840

DB 781 TTTCTGTCTTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 840

QY 841 ATGAGAGAGCAGTATAGATGTTCTCTCGGCTCTCTGAGGGACCTTGGGAGCTCAGGC 900

DB 841 ATGAGAGAGCAGTATAGATGTTCTCTCGGCTCTCTGAGGGACCTTGGGAGCTCAGGC 900

QY 901 TGGGAATCTCCAGGCGAGTAGTCCCTATCAAAATCAAAAGTCCAGTCTTGTGGGGGA 960

DB 901 TGGGAATCTCCAGGCGAGTAGTCCCTATCAAAATCAAAAGTCCAGTCTTGTGGGGGA 960

QY 961 AAACAAAGAGCCCATACCCAGAGGACTGCGGCTTCCCTCACCAGCCAGCTAGGCC 1020

DB 961 AAACAAAGAGCCCATACCCAGAGGACTGCGGCTTCCCTCACCAGCCAGCTAGGCC 1020

QY 1021 TTTCAGAGGAACAAAGACAGACAAATGATTGGCTCTCTGAGGAGATTCAGCCTAG 1080

DB 1021 TTTCAGAGGAACAAAGACAGACAAATGATTGGCTCTCTGAGGAGATTCAGCCTAG 1080

QY 1081 AGCTCTCTCTCCCCCAATCCCTCCCTCGGCTGAGGAACTTAAACAGGAAAAAATTT 1140

DB 1081 AGCTCTCTCTCCCCCAATCCCTCCCTCGGCTGAGGAACTTAAACAGGAAAAAATTT 1140

QY 1141 CGGAAAGCAGATTTAGAGAGCAAAATCCACTGTTGCCCTTGGCTGCCGGGAGAGTG 1200

DB 1141 CGGAAAGCAGATTTAGAGAGCAAAATCCACTGTTGCCCTTGGCTGCCGGGAGAGTG 1200

QY 1201 GACTAGAGAGTCTGGGGCGCAGCCCGAGCCGCTTCCCGCGGCTTTAGGCCGCGG 1260

DB 1201 GACTAGAGAGTCTGGGGCGCAGCCCGAGCCGCTTCCCGCGGCTTTAGGCCGCGG 1260

QY 1261 GGCCCGGCGGGGAGGGGACGACGCCGCGGAGCCCTTAAGACACCTTGTCTACCTTCCA 1320

DB 1261 GGCCCGGCGGGGAGGGGACGACGCCGCGGAGCCCTTAAGACACCTTGTCTACCTTCCA 1320

QY 1321 CCCCAACCCCAACCCACCTTCCCGGAGCTCCCTAGTGTGTTGGGGCGGCTGAACGTGCG 1380

DB 1321 CCCCAACCCCAACCCACCTTCCCGGAGCTCCCTAGTGTGTTGGGGCGGCTGAACGTGCG 1380

QY 1381 CGGTTTAAAGGGGCGGGCGGCTCCAGTGTCTTCTGAGTGAACACTACATATAA 1440

DB 1381 CGGTTTAAAGGGGCGGGCGGCTCCAGTGTCTTCTGAGTGAACACTACATATAA 1440

Db	1381	CGTTTAAAGGCGGCGCCCGCTCCAGTCTCTTCTGAGTGACACTACATAAA	1440			
Qy	1441	CAGAGCCGGGAAACGGCGGGAGAGGAGACACAGGCTTTACCGATAGTAACTC	1500			
Db	1441	CAGAGCCGGGAAACGGCGGGAGGAGGAGACACAGGCTTTGACCGATAGTAACTC	1500			
Qy	1501	TGGCTCGGTGAGCGGAATCTATAAAGGAA	1532			
Db	1501	TGGCTCGGTGAGCGGAATCTATAAAGGAA	1532			
RESULT 3						
AA92831						
Id	AA92831 standard; DNA; 183999 BP.					
XX	AA92831:					
XX	17-MAY-2001 (first entry)					
DT	Human ABC1 genomic DNA.					
XX	High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.					
KW	Homo sapiens.					
OS	WO200115676-A2.					
XX	08-MAR-2001.					
PD	01-SEP-2003; 2000WO-1B01492.					
XX	01-SEP-1999; 9905-0151977.					
PR	15-MAR-2000; 2000US-0526193.					
PR	23-JUN-2000; 2000US-0213958.					
XX	(YBR-) UNIV BRITISH COLUMBIA.					
PA	(XENO-) XENON GENETICS INC.					
XX	Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;					
PI	WPI; 2001-244356/25.					
XX	Treating a lower than normal high density lipoprotein-cholesterol					
PT	(HDL-C) level, a higher than normal triglyceride level, or a					
PT	cardiovascular disease, by administering a compound that modulates LXR-					
PT	or RXR-mediated transcriptional activity -					
XX	Claim 8; Fig 1; 317pp; English.					
PS	The present invention relates to a method for treating a patient					
CC	diagnosed as having a lower than normal high density					
CC	lipoprotein-cholesterol (HDL-C) level, a higher than normal					
CC	triglyceride level, or a cardiovascular disease, involving					
CC	administering a compound that modulates LXR- or RXR-mediated					
CC	transcriptional activity or ABC1 expression or activity.					
CC	The LXR gene product may be used in an assay to identify					
CC	compounds useful for the treatment of a disease or condition selected a					
CC	lower than normal HDL cholesterol level, a higher than normal					
CC	triglyceride level, and a cardiovascular disease.					
XX	Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;					
SQ	Query Match 90.6%; Score 1387.6; DB 22; Length 183999;					
Best Local Similarity 97.7%; Pred. No. 0;						
Matches 1503; Conservative 2; Mismatches 10; Indels 23; Gaps						
Qy	15	GCTCCACATGCACCTCCAGGCGCTGTGG--CCTCTATGGGTCTGCTGAGTGT 73				
Db	27234	GCTCCACATGCACCTCCAGGCGCTGTGGCTTTCTCTATGGGTCTGCTGAGTGT 73				
Qy	74	GATAGAACCACTGATGTAGTACCTGGGCTTGAG--CGTGGCTCGAGATCCTGTGACTG 132				
Db	27294	GATAGAACCACTGATGTAGTACCTGGGCTTGAGCGCGGCTTGAGATCCTGTGACTG 132				

QY	133	TAGCATGGAGGGGCTTGT	-CAGCTGATGCTGTATGACAGTGGTGGGAGTTCCTGGAAT	131
Db	27354	TAGCATGGAGGGGCTTGT	GCAGCTGAATCTGTATGACAGTGGTGGGAGTTCCTGGAAT	27413
QY	192	ATGATGGAGCTGGAGGT	GGGAAGAGAAGTAGGCTGGGGCAGAGCTCTCTATGCAAGCTCA	251
Db	27414	ATGATGGAGCTGGAGGT	GGGAAGAGAAGTAGGCTGGGGCAGAGCTCTCTATGCAAGCTCA	27473
QY	252	TTCTGGCCAAACACT	CAGTGCAAACTGTGAAGAGCTTAAATGTGAATCTGCCCTTCAAGGT	311
Db	27474	TTCTGGCCAAACACT	CAGTGCAAACTGTGAAGAGCTTAAATGTGAATCTGCCCTTCAAGGT	27533
QY	312	GGCTACAAGGTATCT	TGTGCARAGTAGGAGACCTTGGCCCTCCAGTGCACCTTCOCAGG	371
Db	27534	GGCTACAAGGTATCT	TGTGCARAGTAGGAGACCTTGGCCCTCCAGTGCACCTTCOCAGG	27593
QY	372	GCCCTGCTT	-GGCCCTCTTACGGTCTGTCTCAGTCTTCTATGAA--TCTCCCTTCAGG	428
Db	27594	GCCCTGCTT	GGCCCTCTTACGGTCTGTCTCAGTCTTCTATGAA--TCTCCCTTCAGG	27653
QY	429	GGAGATTCATATTT	TAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATAAGGTGACA	488
Db	27654	GGAGATTCATATTT	TAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGATAAGGTGACA	27713
QY	489	TTTAGTTTGTGGCTTGAT	GAACTAAATATTAGACATATGGTGTGTAGCCCTGCA	548
Db	27714	TTTAGTTTGTGGCTTGAT	GAACTAAATATTAGACATATGGTGTGTAGCCCTGCA	27773
QY	549	TTCCACTCTTGCCCTTT	TTTGGCCCTCCAGTGTTTGGGTAGTTTGGTCCCCCTAC	608
Db	27774	TTCCACTCTTGCCCTTT	TTTGGCCCTCCAGTGTTTGGGTAGTTTGGTCCCCCTAC	27832
QY	609	AGCCAAAGCCAAACAGATAAG	CTGGAGCTCTGAGTGCTACATAATTACACGACTGC	668
Db	27833	AGCCAAAGCCAAACAGATAAG	CTGGAGCTCTGAGTGCTACATAATTACACGACTGC	27892
QY	669	AAATCTCTGGCTGCACCTT	CACAAATGTATACAACTAAATACAAAGCTGTGTTTATC	728
Db	27893	AAATCTCTGGCTGCACCTT	CACAAATGTATACAACTAAATACAAAGCTGTGTTTATC	27952
QY	729	ACAGGAGAGCTGATCAATAT	ATAGAAATTAAGAGGGGCTGGTCCCATATTGCTCTGCTGT	788
Db	27953	ACAGGAGAGCTGATCAATAT	ATAGAAATTAAGAGGGGCTGGTCCCATATTGCTCTGCTGT	28011
QY	789	TTTGTGTTGTGTTCTCTTT	-----TTTGTGTTGTGGCTCCCTCTCTCAATT	839
Db	28012	TTTGTGTTGTGTTCTCTTT	-----TTTGTGTTGTGGCTCCCTCTCTCAATT	28071
QY	840	TATGAAGAGACAGTAAAGAT	TTCTCTCGGGTCTCTGAGGGACCTTGGGAGCTCAGG	899
Db	28072	TATGAAGAGACAGTAAAGAT	TTCTCTCGGGTCTCTGAGGGACCTTGGGAGCTCAGG	28131
QY	900	CTGGGAATCTCCAAAGCAGT	AGTGTGCGCTATCAAAATCAAAAGTTCAGGTTTGTGGGGGG	959
Db	28132	CTGGGAATCTCCAAAGCAGT	AGTGTGCGCTATCAAAATCAAAAGTTCAGGTTTGTGGGGGG	28191
QY	960	AAAACAAAAGCAGCCCAT	TACCAAGAGACTGTCCGCTTCCCTCACCACGCTTAGGC	1019
Db	28192	AAAACAAAAGCAGCCCAT	TACCAAGAGACTGTCCGCTTCCCTCACCACGCTTAGGC	28251
QY	1020	CTTTGAAAGAAACAAAGACA	AGCAAAATGATTTGGCGTCTGAGGGAGATTACGCCTA	1079
Db	28252	CTTTGAAAGAAACAAAGACA	AGCAAAATGATTTGGCGTCTGAGGGAGATTACGCCTA	28311
QY	1080	GAGCTCTCTCTCCCCAAAT	CCCTCCCTCCGCGTGAAGAAACTAAACAAAGAAAAAAAT	1139
Db	28312	GAGCTCTCTCTCCCCAAAT	CCCTCCGCTGAGAACTAACAAAGAAAAAAAT	28371
QY	1140	TGCGGAAGCAGGATT	TAGAGGAAGCAAAATTCACCTGTGTCCTTGGCTGCCGGAGAGT	1199
Db	28372	TGCGGAAGCAGGATT	TAGAGGAAGCAAAATTCACCTGTGTCCTTGGCTGCCGGAGAGT	28431


```
FT protein_bind 2051..2059
FT /*tag= ak
FT /bound_moiety= "GATA"
FT 2104..2111
FT /*tag= al
FT /bound_moiety= "SOX5"
FT 2114..2152
FT /*tag= am
FT /bound_moiety= "SYR/HFH/HNF3beta"
FT 2221..2228
FT /*tag= an
FT /bound_moiety= "KZF1"
FT 2234..2249
FT /*tag= ao
FT /bound_moiety= "IK2/NFkappaB/CREL"
FT 2259..2272
FT /*tag= ap
FT /bound_moiety= "LMO2COM/GATA"
FT 2289..2306
FT /*tag= aq
FT /bound_moiety= "KZF1/SRY"
FT 2313..2318
FT /*tag= ar
FT /bound_moiety= "PPAR"
FT 2321..2326
FT /*tag= as
FT /bound_moiety= "PPAR"
FT 2335..2342
FT /*tag= at
FT /bound_moiety= "KZF1"
FT 2361..2384
FT /*tag= au
FT /bound_moiety= "HNF3beta/SRY/EV11"
FT 2426..2433
FT /*tag= av
FT /bound_moiety= "KZF1"
FT 2442..2451
FT /*tag= aw
FT /bound_moiety= "AP4"
FT 2455..2466
FT /*tag= ax
FT /bound_moiety= "SRY"
FT 2491..2498
FT /*tag= ay
FT /bound_moiety= "STAT"
FT 2524..2534
FT /*tag= az
FT /bound_moiety= "STAT/PPAR"
FT 2536..2541
FT /*tag= ba
FT /bound_moiety= "PPAR"
FT 2589..2600
FT /*tag= bb
FT /bound_moiety= "AP2"
FT 2610..2617
FT /*tag= bc
FT /bound_moiety= "KZF1"
FT 2634..2648
FT /*tag= bd
FT /bound_moiety= "LMO2COM/MYOD/E47"
FT 2657..2672
FT /*tag= be
FT /bound_moiety= "RREB1"
FT 2680..2698
FT /*tag= bf
FT /bound_moiety= "KZF1/CMYB"
FT 2728..2740
FT /*tag= bg
FT /bound_moiety= "SPI/GC"
FT 2743..2757
FT /*tag= bh
FT /bound_moiety= "JSE/NMYC/ARNT"
FT 2758..2773
```

```
FT /*tag= bi
FT /bound_moiety= "NFE2A1"
FT 2774..2787
FT /*tag= bj
FT /bound_moiety= "XFD1/HFH"
FT 2794..2806
FT /*tag= bk
FT /bound_moiety= "GC/SP1/KZFI"

Query Match 89.7%; Score 1374.4; DB 24; Length 2910;
Best Local Similarity 97.9%; Prec. No. 0;
Matches 1504; Conservative 0; Mismatches 6; Indels 26; Gaps 10:

Qy 16 GGTCCACATGCACATCCAGGCGCTGTTGG--CTCTCTATGGCTCTGTCTGAGTGT 73
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1345 GCTCCACATGCACATCCAGGCGCTGTTGGCGCTCTCTATGGCTGTCGTCCGAGTGT 1404
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 74 GATAGACCACTGATGTGAGTACCTGGGCTTTCAG--CGTGGGCTGGAGATCCTGTGTGATG 132
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1405 GATAGACCACTGATGTGAGTACCTGGGCTTTCAGCGCTGGGCTGGAGATCCTGTGTGATG 1464
Qy 133 TAGCATGGAGGGGCTTGT--CAGCTGAATGTCTGTATGCAGGTGCTGGAGTCTTGGAAAT 191
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1465 TAGCATGGAGGGGCTTGTGCACTGAATGCTGCACTGCAGGTGTGGAGTCTTGGAAAT 1524
Qy 192 ATGATGGAGCTGGAGTGGGAAGAGAAGTAGCCTTTGGGAGCTCTCTCATGCCACCTCA 251
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1525 ATGATGGAGCTGGAGTGGGAAGAGAAGTAGCCTTTGGGAGCTCTCTCATGCCACCTCA 1584
Qy 252 TTCTGGCCAAAACCTCAGGTCAACACTGTGAAGAGTCTAATATGTGAATCTGCCCTTCAAGGT 311
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1585 TTCTGGCCAAAACCTCAGGTCAACACTGTGAAGAGTCTAATATGTGAATCTGCCCTTCAAGGT 1644
Qy 312 GCCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTTGGGCTCCACGTGCACTTCCAGG 371
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1645 GCCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTTGGGCTCCACGTGCACTTCCAGG 1704
Qy 372 GCTGCTTT--GGCTCTCTTACGGGTCTGCTGAGTCTTCTATGAATCTCCCTTCAGGCG 430
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1705 GCTGCTTTGGGCTCTCTTACGGGTCTGCTGAGTCTTCTATGAATCTCCCTTCAGGCG 1761
Qy 431 AGATTCATATTTAGACTCTTCAACAGTTTGAACCTGAGTTTGGCCAGATAAGTGCACAT 490
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1762 AGATTCATATTTAGACTCTTCAACAGTTTGAACCTGAGTTTGGCCAGATAAGTGCACAT 1821
Qy 491 TAGTTTGTGGCTTGATGAATGACTTAAATATTTAGACATATGGTGTAGCCCTGCATT 550
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1822 TAGTTTGTGGCTTGATGAATGACTTAAATATTTAGAC--ATGGTGTGTAGCCCTGCATT 1879
Qy 551 CCTACTCTTGGCTTTTGTGGGCTTCCAGTGTGTTTGGGTAGTTTGTCTCCCTTACAG 610
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1880 CCTACTCTTGGCTTTTGTGGGCTTCCAGTGTGTTTGGGTAGTTTGTCTCCCTTACAG 1938
Qy 611 CCAAGGCAACAGATAAGTTTGGAGTCTGGAGTGGCTACATAAATTTTACAGACTGCAA 670
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1939 CCAAGGCAACAGATAAGTTTGGAGTCTGGAGTGGCTACATAAATTTTACAGACTGCAA 1998
Qy 671 TTCTCTGGCTGCACCTTCAACAATGTATACAACTAAATAACAAGTCTGTGTTTATATC 730
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1999 TTCTCTGGCTGCACCTTCAACAATGTATACAACTAAATAACAAGTCTGTGTTTATATC 2058
Qy 731 AGGGAGGCTGATCAATATAATGAATTAAGAGGGGCTGCTCCCATATTTGTCTGTGTT 790
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2059 AGGGAGGCTGATCAATATAATGAATTAAGAGGGGCTGCTCCCATATTTGTCTGTGTT 2117
Qy 791 TTG-----TTTGTGTTCTCTTTTGTGTTTGTGGCTCTCTCTCAATTCA 841
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2118 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2177
Qy 842 TGAAGAGAGCAGTAAGATGTTCTCTCGGCTCTCTGAGGAGCTTGGGAGCTCAGGCT 901
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2178 TGAAGAGAGCAGTAAGATGTTCTCTCGGCTCTCTGAGGAGCTTGGGAGCTCAGGCT 2237
```

```
QY 902 GGAATCTCCAGCAGTAGTGCCTATCAAAATCAAGTCAGGTTTGTGGGGGAA 961
DB 2238 GGGAACTCCAAAGCAGTAGTGCCTATCAAAATCAAGTCAGGTTTGTGGGGGAA 2297
QY 962 AACAAAGCAGCCCATACCCAGAGGACITGTCGGCCTTCCCTCACCCAGCCTAGGCCT 1021
DB 2298 AACAAAGCAGCCCATACCCAGAGGACITGTCGGCCTTCCCTCACCCAGCCTAGGCCT 2357
QY 1022 TTGAAGGAAACAAAGACAAAGAAATGATGTGGCGTCTGAGGAGATTCAGCCTAGA 1081
DB 2358 TTGAAGGAAACAAAGACAAAGAAATGATGTGGCGTCTGAGGAGATTCAGCCTAGA 2417
QY 1082 GCTCTCTCCGCCAAATCCCTCCCTCCGCTGAGGAACACTAACAAAGGAAAGAAAAATG 1141
DB 2418 GCTCTCTCCGCCAAATCCCTCCCTCCGCTGAGGAACACTAACAAAGGAAAGAAAAATG 2477
QY 1142 CGGAAGCAGGATTAGAGGAGGAAATTCACCTGCTGCTCCCTTGCTCCGCGGAACGTGG 1201
DB 2478 CGGAAGCAGGATTAGAGGAGGAAATTCACCTGCTGCTCCCTTGCTCCGCGGAACGTGG 2537
QY 1202 ACTAGAGAGTCTGGCGCGCAGCCCGAGCCAGCGCTTCCCGCGCTCTTAGCGCGGCGG 1261
DB 2538 ACTAGAGAGTCTGGCGCGCAGCCCGAGCCAGCGCTTCCCGCGCTCTTAGCGCGGCGG 2597
QY 1262 GCCCGCGCGGGGAGGAGCGCAGACCGCGGACCTTAAGACACCTGTGTACCCCTCCAC 1321
DB 2598 GCCCGCGCGGGGAGGAGCGCAGACCGCGGACCTTAAGACACCTGTGTACCCCTCCAC 2657
QY 1322 -----CCCCACCCACCCCTCCCGCAACTCCCTAGATGTGTCGTGGCGGCTGAAGC 1376
DB 2658 CCCACCCACCCACCCCTCCCGCAACTCCCTAGATGTGTCGTGGCGGCTGAAGC 2717
QY 1377 TGGCCCGCTTAAGGGCGGGGGCGGGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436
DB 2718 TGGCCCGCTTAAGGGCGGGGGCGGGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2777
QY 1437 TAAACAGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1496
DB 2778 TAAACAGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2837
QY 1497 COTCTGCGCTCGGTGCAGCCGAATCTATAAAGGAA 1532
DB 2838 COTCTGCGCTCGGTGCAGCCGAATCTATAAAGGAA 2873
```

```
RESULT 5
AAD37265
ID AAD37265 standard; DNA; 3231 bp.
AC AAD37265;
XX
XX
DT 21-AUG-2002 (first entry)
XX
XX Human ABC1 transcription regulatory DNA #1.
DE
XX
XX Human: ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200183746-A2.
XX
XX 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-EP05486.
XX
XX 02-MAY-2000; 2000US-201280P.
XX
XX (AVET ) AVENTIS PHARMA SA.
XX
XX Rosier-Montus M, Prades C, Lemoine C, Naudin L, Benefile P;
PI Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
XX
```

```
DR WPI; 2002-154404/20.
XX
XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
PT (ABC1) and screening for candidate modulatory compounds or substances
PT
XX
XX Claim 1; Page 130-131; 152pp; English.
XX
XX The invention relates a nucleic acid which is capable of regulating the
CC transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
CC casual gene for pathologies linked to a dysfunctioning of cholesterol
CC metabolism, including diseases such as atherosclerosis. Polynucleotides
CC that are capable of modulating the transcription of the ABC1 gene. They
CC are used in antisense therapy. Compositions comprising sequences of the
CC invention are used to treat hypercholesterolaemia and atherosclerosis.
CC The present sequence is human ABC1 transcription regulating DNA.
XX
XX Sequence 3231 bp; 809 A; 773 C; 876 G; 773 T; 0 other;
```

```
Query Match 89.7%; Score 1374.4; DB 24; Length 3231;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

QY 16 GGCTCCACATGCACATTCACAGGCCCTGCTGG--CTCTTCTATGGTCTGTCCAGTGT 73
DB 1345 GCCTCACAIGACATTCACAGGCCCTGCTGGCCCTCTTCTATGCCCTGTCTCTGAGTG 1404
QY 74 GATACAACACATGATGATGATACCTGGGCTTGAG--CGTGGCCTGGAGATCCTGTGACTG 132
DB 1405 GATACAACACATGATGATGATACCTGGGCTTGAGCCGTGGCCTGGAGATCCTGTGACTG 1464
QY 133 TAGCATGGAGGGGGCTTGT--CAGCTGAATGTCTGTATCAGGTGGTGGAGTCTCTGGAAT 191
DB 1465 TAGCATGGAGGGGGCTTGTGCGAGCTGAATGTCTGCATGCAGGTGGTGGAGTCTCTGGAAT 1524
QY 192 ATGATGGAGCTGGAGTGGGAGAGAGTAGCTTGGGGCAGCTCTCTCATGCCACCTCA 251
DB 1525 ATGATGGAGCTGGAGTGGGAGAGAGTAGCTTGGGGCAGCTCTCTCATGCCACCTCA 1584
QY 252 TCTGGCCAAAACCTCAGGTCAAACGTGAAGAGTCTAAATGTGAATCTGCCCTTCAAGGT 311
DB 1585 TCTGGCCAAAACCTCAGGTCAAACGTGAAGAGTCTAAATGTGAATCTGCCCTTCAAGGT 1644
QY 312 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTTGTGGCCTCCACGTGCACTTCCAGG 371
DB 1645 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTTGTGGCCTCCACGTGCACTTCCAGG 1704
QY 372 GGCTGCT--GGCCTCTCTACGGGTCTGTCTCCTGAGTCTTCTATGAATCTCCCTTCAGGC 430
DB 1705 GGCTGCTGGGCTCTCTACGGGTCTGTCTCCTGAGTCTTCTATGAATCTCCCTTCAGGC 1761
QY 431 AGATTCAATATTAGACTCTTTCACAGTTTGACCTGAGTTTGGCCAGAAATAAGTGCACATT 490
DB 1762 AGATTCAATATTAGACTCTTTCACAGTTTGACCTGAGTTTGGCCAGAAATAAGTGCACATT 1821
QY 491 TASTTTGTGGCTTGCATGAATGACTTAATATTAGACATATGGTGTGAGCCCTGCATT 550
DB 1822 TAGTTTGTGGCTTGCATGAATGACTTAATATTAGAC--ATGGTGTGTAGSCCTGCATT 1879
QY 551 CCTACTCTTGGCTTTTCTTTTGGCCCTCCAGTCTTGGGTAGTTTGGTCCCTCCCTACAG 610
DB 1880 CCTACTCTTGGCTTTTCTTTTGGCCCTCCAGTCTTGGGTAGTTTGGT--CCCTACAG 1938
QY 611 CCAAGGCAACACAGATAAGITGGAGGTCTGGAGTGGCTGACATTAATTTACAGCACTGCAA 670
DB 1939 CCAAGGCAACACAGATAAGITGGAGTGGCTGACATTAATTTACAGCACTGCAA 1996
QY 671 TTCTCTGGCTGCACCTTCACAAATGTATACAAACTAAATACAAAGTCTGTGTTTATCAC 730
DB 1999 TTCTCTGGCTGCACCTTCACAAATGTATACAAACTAAATACAAAGTCTGTGTTTATCAC 2058
QY 731 AGGGAGGCTGATCAATATAATGAATTAAGGGGGGCTGGTCCCATATTTGTTCTGTGTT 790
```



```

QY 940 AGTCCAGGTTTGTGGGGGAAACAAACAGCAGCCATACCCAGAGGAGTGTCCGCCTT 999
Db 479 AAGTCCAGGTTTGTGGGGGAAACAAACAGCAGCCATACCCAGAGGAGTGTCCGCCTT 538
QY 1000 CCCTCACCCTCAGCCTAGGCTTTGAAAGGAAACAAAAGACAAAGCAAAATGATTGGCGT 1059
Db 539 CCCTCACCCTCAGCCTAGGCTTTGAAAGGAAACAAAAGACAAAGCAAAATGATTGGCGT 598
QY 1060 CCGAGGAGATTACGCTTAGAGCTCTCTCTCCGCCAATCCCTCCCTCCGCTGAGGAAA 1119
Db 599 CCGAGGAGATTACGCTTAGAGCTCTCTCT-CCCCAATCCCTCCCTCCGCTGAGGAAA 657
QY 1120 CTAACAAGCAAAATAAATTCGGGAAGCAGGATTTAGAGGAGCAAAATCCCACTGGTG 1179
Db 558 CTAACAAGCAAAATAAATTCGGGAAGCAGGATTTAGAGGAGCAAAATCCCACTGGTG 717
QY 1180 CCCTTGCTCCCGGACGTGGACTAGAGAGTCTGGGGGCGAGCCCGAGCCCGAGCGCTT 1239
Db 718 CCCTTGCTCCCGGACGTGGACTAGAGAGTCTGGGGGCGAGCCCGAGCCCGAGCGCTT 777
QY 1240 CCGCGCGCTTTAGCGCGGCGCGCGCGGCGGAGAGGGAGCGACACCGCGGACCCCTA 1299
Db 778 CCGCGCGCTTTAGCGCGGCGCGCGCGGCGGAGAGGGAGCGACACCGCGGACCCCTA 837
QY 1300 AGACACCTGCTGTACCTCCACCCCGACCCACCCACCTCCCGCCACTCCCTAGAGTGTG 1359
Db 838 AGACACCTGCTGTACCTCCACCCCGACCCACCCACCTCCCGCCACTCCCTAGAGTGTG 897
QY 1360 TCGTGGCGGCTGAACCTCCCGCTTTAAGGGGGCGGCGCGCGCTCCACGCTGCTTTCTGC 1419
Db 898 TCGTGGCGGCTGAACCTCCCGCTTTAAGGGGGCGGCGCGCGCTCCACGCTGCTTTCTGC 957
QY 1420 IGAGTACTGAATACATACAGAGCCCGGACCGGGGCGGGGAGGAGGAGGAGGACAG 1479
Db 958 TGAGTACTGAATACATACAGAGCCCGGACCGGGGCGGGGAGGAGGAGGAGGAGGACAG 1017
QY 1480 GCTTTGACCGATAGTAACCTCTCGGCTCGGTGAGCGCGAATCTATAAAGGAA 1532
Db 1018 GCTTTGACCGATAGTAACCTCTCGGCTCGGTGAGCGCGAATCTATAAAGGAA 1070

RESULT 7
AAD37267/c
ID AAD37267 standard; DNA; 2910 BP.
XX AC AAD37267;
XX XX
XX 21-AUG-2002 (first entry)
XX Human ABC1 transcription regulatory DNA #3.
XX Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX OS Homo sapiens.
XX FH
XX Key Location/Qualifiers
FT protein_bind 12..23 /*tag= a
FT /*bound_moiety= "LMO2COM/MYOD"
FT protein_bind 97..107 /*tag= b
FT /*bound_moiety= "DeltaEFl"
FT protein_bind 110..125 /*tag= c
FT /*bound_moiety= "S8/NKX2.5"
FT protein_bind 196..211 /*tag= d
FT /*bound_moiety= "S8"
FT protein_bind 228..237 /*tag= e
FT /*bound_moiety= "GATA"

protein_bind 399..410 /*tag= f
/*bound_moiety= "IK2"
protein_bind 412..420 /*tag= g
/*bound_moiety= "LYF1"
protein_bind 528..539 /*tag= h
/*bound_moiety= "LMO2COM/MYOD/DeltaEFl"
protein_bind 549..556 /*tag= i
/*bound_moiety= "LYF1"
protein_bind 558..568 /*tag= j
/*bound_moiety= "DeltaEFl"
protein_bind 590..596 /*tag= k
/*bound_moiety= "LKX2.5"
protein_bind 608..620 /*tag= l
/*bound_moiety= "NFY/CAT"
protein_bind 708..715 /*tag= m
/*bound_moiety= "M2Fl"
protein_bind 723..730 /*tag= n
/*bound_moiety= "M2Fl"
protein_bind 771..785 /*tag= o
/*bound_moiety= "HFH2/SRY/EV11"
protein_bind 803..812 /*tag= p
/*bound_moiety= "CREBP1/VBP"
protein_bind 831..837 /*tag= r
/*bound_moiety= "NKX2.5"
protein_bind 1076..1089 /*tag= s
/*bound_moiety= "GATA"
protein_bind 1173..1188 /*tag= t
/*bound_moiety= "LXralpha/deltaEFl"
protein_bind 1189..1209 /*tag= u
/*bound_moiety= "DeltaEFl/LYF1/IK2"
protein_bind 1483..1491 /*tag= v
/*bound_moiety= "AP4"
protein_bind 1498..1514 /*tag= w
/*bound_moiety= "LMO2-COM/MYOD/deltaEFl/E47"
protein_bind 1524..1545 /*tag= x
/*bound_moiety= "ZIP/deltaEFl"
protein_bind 1597..1607 /*tag= y
/*bound_moiety= "DeltaEFl"
protein_bind 1622..1627 /*tag= z
/*bound_moiety= "PPAR"
protein_bind 1632..1637 /*tag= aa
/*bound_moiety= "PPAR"
protein_bind 1685..1698 /*tag= ab
/*bound_moiety= "USF/NMYC/MYOMAX"
protein_bind 1787..1797 /*tag= ac
/*bound_moiety= "DeltaEFl"
protein_bind 1809..1819 /*tag= ad
/*bound_moiety= "DeltaEFl"
protein_bind 1822..1833

```

```

FT      /tag= ae
FT      /bound_moiety= "SRY"
FT      /tag= af
FT      /bound_moiety= "AP1"
FT      /tag= ag
FT      /bound_moiety= "HNF3beta"
FT      /tag= ah
FT      /bound_moiety= "NKX2.5"
FT      /tag= ai
FT      /bound_moiety= "PPAR/NKX2.5/PPAR"
FT      /tag= aj
FT      /bound_moiety= "PPAR/NKX2.5/PPAR"
FT      /tag= ak
FT      /bound_moiety= "GATA"
FT      /tag= al
FT      /bound_moiety= "SOX5"
FT      /tag= am
FT      /bound_moiety= "SYR/HFH/HNF3beta"
FT      /tag= an
FT      /bound_moiety= "MZFL"
FT      /tag= ao
FT      /bound_moiety= "IK2/NFkappaB/CREL"
FT      /tag= ap
FT      /bound_moiety= "LMO2/COM/CAAT"
FT      /tag= aq
FT      /bound_moiety= "MZFL/SRY"
FT      /tag= ar
FT      /bound_moiety= "PPAR"
FT      /tag= as
FT      /bound_moiety= "PPAR"
FT      /tag= at
FT      /bound_moiety= "MZFL"
FT      /tag= au
FT      /bound_moiety= "HNF3beta/SRY/EVIL"
FT      /tag= av
FT      /bound_moiety= "MZFL"
FT      /tag= aw
FT      /bound_moiety= "AP4"
FT      /tag= ax
FT      /bound_moiety= "SRY"
FT      /tag= ay
FT      /bound_moiety= "STAT"
FT      /tag= az
FT      /bound_moiety= "STAT/PPAR"
FT      /tag= ba
FT      /bound_moiety= "PPAR"
FT      /tag= bb
FT      /bound_moiety= "AP2"
FT      /tag= bc

```

```

FT      /bound_moiety= "MZFL"
FT      /tag= bd
FT      /bound_moiety= "LMO2/COM/XYOD/347"
FT      /tag= be
FT      /bound_moiety= "RREB1"
FT      /tag= bf
FT      /bound_moiety= "MZFL/CMYB"
FT      /tag= bg
FT      /bound_moiety= "SP1/GC"
FT      /tag= bh
FT      /bound_moiety= "USF/NMYC/ARNT"
FT      /tag= bi
FT      /bound_moiety= "NFE2A1"
FT      /tag= bj
FT      /bound_moiety= "XFD1/HFH"
FT      /tag= bk
FT      /bound_moiety= "GC/SP1/MZF1"

Query Match      6.7%; Score 102.2; DS 24; Length 2910;
Best Local Similarity 55.9%; Pred. No. 1.5e-17;
Matches 164; Conservative 0; Mismatches 83; Indels 2; Gaps 1;

QY 241 ATCCACCTCATTCCTGGCCAAACTCAGTCAAACTGACAGAGTCTAAATCTGAATCTG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1820 ATGTCACCTTATTCCTGGCCAAACTCAGTCAAACTGACAGAGTCTAAATCTGAATCTG 1761

QY 301 CCCTTCAGAGTGGCTACAAAGGTATCTTTGTCAAGTAGGAGACCTTGWGGCCTCCACGT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1760 CCTGAAGGATTCATAGAGACTCAGGACAGACCCCTGAGAGGCCCAAGCAGGCCCTG 1701

QY 361 GCACITCCAGGCCCTGTGGCCTCTCTCTACGGGCTCTCTGAGTCTTCTATGATCTC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1700 GAAGTGCACCTGGAGGC--CACAAAGGTCTCTTACCTTGACAAGATACCTTTGTAGCCAC 1643

QY 421 CCTTCAGGCGCAGATTATATTATAGACTCTTCACAGTTTGCAGCTGAGTTTGGCCAGAATA 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1642 CTTGAAGGCGCAGATTACATTTAGACTCTTCACAGTTTGCAGCTGAGTTTGGCCAGAATA 1583

QY 481 AGGTGACAT 489
    ||| ||| |||
DB 1582 AGGTGGCAT 1574

RESULT 8
AAD37265/c
ID AAD37265 standard; DNA; 3231 BP.
XX
AC AAD37265;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #1.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200183746-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-EP05488.
XX
PR 02-MAY-2000; 2000US-201280P.

```

XX (AVEI) AVENTIS PHARMA SA.
 PA Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
 PI Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
 XX WPI; 2002-154404/20.
 DR
 XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
 PT (ABC1) and screening for candidate modulatory compounds or substances
 PT
 XX Claim 1; Page 130-131; 152pp; English.
 PS
 XX The invention relates a nucleic acid which is capable of regulating the
 CC transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
 CC casual gene for pathologies linked to a dysfunctioning of cholesterol
 CC metabolism, including diseases such as atherosclerosis. Polynucleotides
 CC of the invention are used to screen candidate molecules or substances
 CC that are capable of modulating the transcription of the ABC1 gene. They
 CC are used in antisense therapy. Compositions comprising sequences of the
 CC invention are used to treat hypercholesterolemia and atherosclerosis.
 CC The present sequence is human ABC1 transcription regulating DNA.
 XX
 SQ Sequence 3231 BP; 809 A; 773 C; 876 G; 773 T; 0 other;
 Query Match 6.7%; Score 102.2; DB 24; Length 3231;
 Best Local Similarity 65.9%; Pred. No. 1.6e-17;
 Matches 164; Conservative 0; Mismatches 83; Indels 2; Gaps 1;
 QY 241 ATGCCACCTCATTCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTAATCTG 300
 ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1820 ATGTCACCTTATTCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTAATCTG 1761
 QY 301 CCCTTCAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGGCCCTCCAGG 360
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1760 CCCTGAGGAGTTCATAAGAGACICAGGACAGACCCGTAGAAGGCCCAAGCAGGCCCTG 1701
 QY 361 GCACCTTCAGGCGCTGCTGGCCCTCTTCTACGGGTCTCTCTGAGTCTCTGATGATGAATCTC 420
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1700 GAAGTGCAGTGGAGGC--CACAGGTCTCTACCTTGACAAAGATACCTTTGTAGCCAC 1643
 QY 421 CCTTCAGGCGCAGATTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAATA 480
 ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1642 CTTGAAGGCGCAGATTTCACATTTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAATG 1583
 QY 481 AGGTGACAT 489
 |||| ||||
 Db 1582 AGGTGGCAT 1574
 RESULT 9
 AAF24681/c
 ID AAF24681 standard; DNA; 1643 BP.
 XX
 AC AAF24681;
 XX
 XX 20-APR-2001 (first entry)
 DT
 XX Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
 XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200078972-A2.
 PN
 XX 28-DEC-2000.
 PD
 XX 16-JUN-2000; 2000WO-US16765.
 PF

XX 18-JUN-1999; 99US-0140264.
 PR 14-SEP-1999; 99US-0153872.
 PR 19-NOV-1999; 99US-0166573.
 XX
 PA (CVTH-) CV THERAPEUTICS INC.
 XX
 XX Lawn RM, Wade D, Garvin M;
 PI WPI; 2001-137812/14.
 DR
 XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
 PT useful for the development of agents for the treatment of heart disease
 PT and other disorders associated with hypercholesterolemia and
 PT atherosclerosis -
 XX
 PS Claim 1; Page 143-144; 215pp; English.
 XX
 CC The present sequence represents the 5' flanking region of the human
 CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
 CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
 CC variety of substrates across the plasma membrane. ABC1 is a pivotal
 CC protein in the apolipoprotein-mediated mobilisation of intracellular
 CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
 CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
 CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
 CC are useful for developing pharmaceutical agents for the treatment of
 CC heart disease and other disorders associated with hypercholesterolemia
 CC and atherosclerosis. The genes are useful for developing screening assays
 CC to screen for compounds that regulate the expression of genes associated
 CC with cholesterol transport. The genes and proteins are also useful for
 CC are also useful as diagnostic indicators of cardiovascular disease and
 CC other disorders associated with hypercholesterolemia.
 XX
 SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;
 Query Match 6.3%; Score 97; DB 22; Length 1643;
 Best Local Similarity 65.8%; Pred. No. 3.2e-16;
 Matches 154; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 241 ATGCCACCTCATTCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTAATCTG 300
 ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 489 ATGTCACCTTATTCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTAATCTG 430
 QY 301 CCCTTCAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGGCCCTCCAGG 360
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 429 CCCTGAGGAGATTCATAGAGACTCAGGACAGACCCGTAGAAGGCCCAAGCAGGCC 370
 QY 361 GCACCTTCAGGCGCTGCTGGCCCTCTTCTACGGGTCTCTCTGAGTCTCTGATGATGATCTC 420
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 369 TGAAGTGCACCTGGAGGCCACAAAGGTCTCTCTACCTTGACAAAGATACCTTTGTAGCCAC 310
 QY 421 CCTTCAGGCGCAGATTCATATTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAATA 480
 ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 309 CTTGAAGGCGCAGATTTCACATTTAGACTCTTCACAGTTTGACCTGAGTTTGGCCAGAATG 250
 QY 481 AGGTGACAT 489
 |||| ||||
 Db 249 AGGTGGCAT 241
 RESULT 10
 AAF24703/c
 ID AAF24703 standard; DNA; 1643 BP.
 XX
 AC AAF24703;
 XX
 XX 20-APR-2001 (first entry)
 DT
 XX Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
 XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
 KW atherosclerosis; cholesterol transport; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200078972-A2.
 PN
 XX 28-DEC-2000.
 PD
 XX 16-JUN-2000; 2000WO-US16765.
 PF

KW chomosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

PN WO200078971-A2.

PD 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16591.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

PA (UNIW) UNIV WASHINGTON.

PI Lawn RM, Wade D, Oram JF, Garvin M;

DR WPI; 2001-137811/14.

XX Adenosine triphosphate (ATP) binding cassette protein (ABC) i

PT polynucleotides and polypeptides, useful for treatment of heart disease

PT and other disorders associated with hypercholesterolemia and

PT atherosclerosis -

XX Disclosure; Page 138-139; 211pp; English.

XX The present sequence represents the 5' flanking region of the human

CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1

CC resides in cell membranes and utilizes ATP hydrolysis to transport a wide

CC variety of substrates across the plasma membrane. ABC1 is a pivotal

CC protein in the apolipoprotein-mediated mobilisation of intracellular

CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic

CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1

CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins

CC are useful for developing pharmaceutical agents for the treatment of

CC heart disease and other disorders associated with hypercholesterolemia

CC and atherosclerosis. The genes are useful for developing screening assays

CC to screen for compounds that regulate the expression of genes associated

CC with cholesterol transport. The genes and proteins are also useful for

CC are also useful as diagnostic indicators of cardiovascular disease and

CC other disorders associated with hypercholesterolemia.

XX Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 6.3%; Score 97; DB 22; Length 1643;

Best Local Similarity 51.8%; Pred. No. 3.2e-16;

Matches 154; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 241 ATGCCACCTCATCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 300

DB 489 ATGTACCTTAATCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 430

QY 301 CCTTCAAGTGTGTACAAAGTATCTTTGTCAGAGTGTAGAGACCTTGTGGCTCCACGT 360

DB 429 CCTTGAAGGAGATTCATAGAGACTCAGGACAGACCCGTAGAAGAGGCGCCAGCGCC 370

QY 361 GCACCTCCAGGGCTCTGTGGCTCTTCTACGGGTCTGTGCTGTCTTCTATGAATCTC 420

DB 369 TGAAGTGCACGTGGAGGCGCACAAAGTCTCTTACCTTGACAAAGATACCTTTGTAGCCAC 310

QY 421 CCTTCAGGCGAGATTCATATTTAGACTCTTCACAGTTTACCTTGTGGTGGCCAGAAATA 480

DB 309 CTGAAGGCGAGATTCATATTTAGACTCTTCACAGTTTACCTTGTGGTGGCCAGAAATA 250

QY 481 AGGTGACAT 489

DB 249 AGGTGGCAT 241

RESULT 11

AAF92831/C
ID AAF92831 standard; DNA; 183999 BP.

XX AC AAF92831;

XX 17-MAY-2001 (first entry)

XX Human ABC1 genomic DNA.

XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

XX Homo sapiens.

PN WO200115676-A2.

XX 08-MAR-2001.

XX 01-SEP-2000; 2000WO-IB01492.

XX 01-SEP-1999; 99US-0151977.

XX 15-MAR-2000; 2000US-0526193.

XX 23-JUN-2000; 2000US-0213958.

XX (UYBR-) UNIV BRITISH COLUMBIA.

PA (XENO-) XENON GENETICS INC.

PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;

XX WPI; 2001-244356/25.

XX Treating a lower than normal high density lipoprotein-cholesterol

CC (HDL-C) level, a higher than normal triglyceride level, or a

CC cardiovascular disease, by administering a compound that modulates LXR-

CC or RXR-mediated transcriptional activity -

XX Claim 8; Fig 1; 317pp; English.

XX The present invention relates to a method for treating a patient

CC diagnosed as having a lower than normal high density

CC lipoprotein-cholesterol (HDL-C) level, a higher than normal

CC triglyceride level, or a cardiovascular disease, involving

CC administering a compound that modulates LXR- or RXR-mediated

CC transcriptional activity or ABC1 expression or activity.

CC The LXR gene product may be used in an assay to identify

CC compounds useful for the treatment of a disease or condition selected a

CC lower than normal HDL cholesterol level, a higher than normal

CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 6.0%; Score 92.6; DB 22; Length 183999;

Best Local Similarity 63.1%; Pred. No. 5.4e-14;

Matches 159; Conservative 0; Mismatches 90; Indels 3; Gaps 1;

QY 241 ATGCCACCTCATCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 300

DB 27714 ATGTACCTTAATCTGGCCAAACTCAGTCAAACTGTGAAGAGTCTAAATGTGAATCTG 27655

QY 301 CCTTCAAGTGTGTACAAAGTATCTTTGTCAGAGTGTAGAGACCTTGTGGCTCCACGT 360

DB 27654 CCTTGAAGGAGATTCATAGAGACTCAGGACAGACCCGTAGAAGAGGCGCCAGCAGG 27595

QY 361 GCACCTCCAGGGCTCTGTGGCTCTTCTACGGGTCTGTGCTGTCTTCTATGAAT 417

DB 27594 CCTTGAAGTGCACGTGGAGGCGCACAAAGTCTCTTACCTTGACAAAGATACCTTTGTAGC 27535

QY 418 CTCCTTCAGGCGAGATTCATATTTAGACTCTTTCACAGTTTACCTTGTGGTGGCCAGA 477

DB 27534 CACCTTGAAGGCGAGATTCACATTTAGACTCTTTCACAGTTTACCTTGTGGTGGCCAGA 27475

QY 478 ATAGGTGACAT 489

DB 27474 ATGAGGTGGCAT 27463

Qy 816 TTTGTGGCTCCTTCTCTCAATTTATGAAGAGAGAGTAAG 858
||| | | | | | | | | | | | | | | | | | | | | |
Db 49 TTTTGTGGTCAGAGCGGTCAAGTTAAGTTGAAATCICCTAAG 7

Search completed: February 3, 2003, 16:28:38
Job time : 933.553 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:35:45 ; Search time 2898.81 Seconds
(without alignments)
8559.210 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1_1532

Perfect score: 1532
Sequence: 1 gaatctctgtggtcgctc.....agccgaatctataaaaggaa 1532

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmd:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_btc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	245.4	16.0	289	10	AW748338
2	111.4	7.3	467	9	AA527406
3	60.8	4.0	1201	17	CNS015W3
4	54.8	3.6	298	17	CNS04M2X
5	53.8	3.5	844	17	CNS0052P
6	53	3.5	237	9	AU222489

C	7	52.4	3.4	910	17	CNS0060N
C	8	51.2	3.3	832	17	CNS015XC
C	9	51	3.3	681	17	CNS02FE9
C	10	50.6	3.3	925	17	CNS0091P
C	11	50.4	3.3	844	17	CNS0052P
C	12	50.4	3.3	870	17	CNS005E2
C	13	48.4	3.2	1004	14	BQ944204
C	14	48	3.1	900	17	AG081217
C	15	47.6	3.1	1009	17	CNS010FW
C	16	47.4	3.1	1059	14	BQ680479
C	17	47.4	3.1	1101	17	CNS017V2
C	18	47.2	3.1	477	9	AL513813
C	19	47.2	3.1	925	17	CNS0091P
C	20	47.2	3.1	942	17	AG042878
C	21	47	3.1	658	12	BF968907
C	22	47	3.1	843	17	CNS00S1
C	23	47	3.1	973	17	CNS008OF
C	24	47	3.1	997	17	CNS006DN
C	25	46.8	3.1	431	9	AL513947
C	26	46.8	3.1	976	17	CNS0068V
C	27	46.6	3.0	978	17	AG030617
C	28	46.6	3.0	1101	17	CNS0083U
C	29	46.2	3.0	527	9	AL514325
C	30	46.2	3.0	875	17	AG043475
C	31	46.2	3.0	1136	14	BQ943816
C	32	46.2	3.0	1417	14	BM610236
C	33	46	3.0	758	17	AG073529
C	34	45.8	3.0	203	10	AW150511
C	35	45.8	3.0	471	10	BE503803
C	36	45.8	3.0	523	9	AL514015
C	37	45.8	3.0	766	17	AG041031
C	38	45.8	3.0	901	17	CNS0760F
C	39	45.8	3.0	932	17	CNS00720
C	40	45.8	3.0	987	17	CNS017LB
C	41	45.8	3.0	997	14	BQ953024
C	42	45.6	3.0	223	13	BG995916
C	43	45.6	3.0	235	10	BE012563
C	44	45.6	3.0	456	9	AL554951
C	45	45.6	3.0	526	9	AL657934

ALIGNMENTS

RESULT 1
AW748338/C
LOCUS
DEFINITION RC6-BT0252-271099-012-cl0 BT0252 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW748338
VERSION AW748338.1 GI:7663270
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 289)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunslein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7). 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-BT0252-271099-012-cl08t3=1999-10-27&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 146
 High quality sequence stop: 289.

FEATURES

Location/Qualifiers
 1..289
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT0252"
 /dev_stage="Adult"

/note="Organ: breast; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (J.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 94 a 70 c 51 g 74 t

ORIGIN

Query Match 15.0%; Score 245.4; DB 10; Length 289;
 Best Local Similarity 96.9%; Pred. No. 1.6e-38;
 Matches 282; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 678 GTGCACCTTCACAAAGTATACAACTAAATACAAAGTCCTGTTTATATACACAGGAGG 737

|||||

DB 289 GTGCACCTTCACAAAGTATACAACTAAATACAAAGTCCTGTTTATATACACAGGAGG 230

|||||

QY 738 CTGATCAATATATGAATATAAAGGGGCGTGGTCCCATATTTCTGTGTTTGTGTTG 797

|||||

DB 229 CTGATCAATATATGAATATAAAGGGGCGTGGTCCCATATTTCTGTGTTTGTGTTG 171

|||||

QY 798 TTGTTTCTTTTGTGTTTGTGCGCTCTCTCTCAATTTATGAAGAGAGCAGTAA 857

|||||

DB 170 TTGTTTCTTTTGTGTTTGTGCGCTCTCTCTCAATTTATGAAGAGAGCAGTAA 112

|||||

QY 858 GATGTCCTCTCGGGTCCTCTCAGGAGACCTGGGAGCTCAGGCTGGGAATCTCCAAGGCA 917

|||||

DB 111 GATGTCCTCTCGGGTCCTCTCAGGAGACCTGGGAGCTCAGGCTGGGAATCTCCAAGGCA 53

|||||

QY 918 GTAGGTCGCTATCAAAATCAAGTCCAGGTTTGCGGGGAAACAAAA 968

|||||

DB 52 GTAGGTCGCTATCAAAATCAAGTCCAGGTTTGCGGGGAAACAAAA 2

|||||

RESULT 2
 LOCUS AA527406 467 bp mRNA linear EST 21-AUG-1997
 DEFINITION ng37c02.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:935962 3', mRNA sequence.

ACCESSION AA527406

VERSION AA527406.1 GI:2269475

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 467)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov

Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.,
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.lnl.gov/bbrp/image/image.html
 Insert Length: 1057 Std Error: 0.00
 Seq primer: -40ml3 fwd. EI from Amersham
 High quality sequence stop: 380.

FEATURES

Location/Qualifiers
 1..467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:935962"
 /clone_lib="NCI_CGAP_Co3"
 /sex="pooled"
 /tissue_type="colon"

/lab_host="DH10B"
 /note="vector: p7730-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization."

BASE COUNT 72 a 118 c 166 g 111 t

ORIGIN

Query Match 7.3%; Score 111.4; DB 9; Length 467;

Best Local Similarity 99.1%; Pred. No. 4.4e-12;

Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 816 TTGTGGCTCTCTCTCAATTTATGAAGAGAGCAGTAAAGTTCCTCTCGGGTCC 875

|||||

DB 1 TTGTGGCTCTCTCTCAATTTATGAAGAGAGCAGTAAAGTTCCTCTCGGGTCC 60

|||||

QY 876 TCTGAGGACCTGGGAGCTCAGGCTGGGAATCTCCAAGGAGTAGTCGCT 928

|||||

DB 61 TCTGAGGACCTGGGAGCTCAGGCTGGGAATCTCCAAGGAGTAGTCGCT 113

|||||

RESULT 3

CNS015W3

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence SP6 end of BAC

BACN15123 of DrosBAC library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

ACCESSION AL105981

VERSION AL105981.1 GI:5619397

KEYWORDS GSS.

SOURCE Drosophila melanogaster.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1201)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Genoscope.

COMMENT Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ;

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the European Drosophila Genome Project (EDGP) -

http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC

library (Dros-BAC) was made by Alain Billaud at CEPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton

and Genevieve Payan. It has been constructed in the vector

pBeloeBAC11.

FEATURES

Location/Qualifiers

1..1201

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="BACN15123"

/clone_lib="DrosBAC"


```

: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZ9pt-F1s
: US-08-232-463-14

Query Match          2.7%  Score 40.6; DB 1; Length 7218;
Best Local Similarity 4.2%  pred.No. 0.34; Mismatches 125; Indels 0; Gaps 0;
Matches 13; Conservative 171;

QY 298 CTGCCCTTCAAGTGGCTGCTCAAGGATATCTTGTCAAGGTAGGAGACCTTGTGGCTTCA 357
    |||| | ||| ||| | : : : : : : : : : : : : : : : : : : : : : :
Db 1043 CTGCAGGTGAGGAGCTGCGATYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1102

QY 358 CGTGCACCTTCAGGCGCTGCTGGCTCTTCIAGGGTCTCTCGAGTCTCTTAAGAT 417
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1203 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1162

QY 418 CTCCTTCAGGCGAGATCATATATAGACTCTTCACAGTTTGACTGAGTTTGGCCAGA 477
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1163 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1222

QY 478 ATAGGTGACATTTAGTTTGTGGCTTGATGAATGACTTAATAATTAGACATATGGTGT 537
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1223 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1282

QY 538 TAGGCTGCTACTCTGCTACTCTGCTCTTTTGTGGCTCTCCAGTGTTCGGGTAGTTT 597
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1283 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1342

QY 598 GCTCCGCT 606
    : : : : : :
Db 1343 YVYVYVYVYV 1351

RESULT 3
US-08-483-533-4
: Sequence 4, Application US/08483533
: Patent No. 6172047
: GENERAL INFORMATION:
: APPLICANT: Roizman, Bernard
: APPLICANT: Chou, Joany
: TITLE OF INVENTION: Method For Treating Tumorigenic
: TITLE OF INVENTION: Diseases
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/483,533
: FILING DATE: 07-MAR-95
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/419,853
: FILING DATE: 11-APR-95
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/861,233
: FILING DATE: 31-MAR-92
: ATTORNEY/AGENT INFORMATION:
: NAME: Zeller, James P.
: REGISTRATION NUMBER: 28,491
: REFERENCE/DOCKET NUMBER: 28097/32742
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 595 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-483-533-4

Query Match          2.5%  Score 38.8; DB 4; Length 595;
Best Local Similarity 49.2%  pred.No. 0.33; Mismatches 132; Indels 3; Gaps 1;
Matches 131; Conservative 0;

QY 1171 CCACCTGTGCTCCCTTGGTGTCCGGGAACGTGACATAGAGTCTCGGGCGACGCCGAGC 1230
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 99 CCGCAGACACCTCGCGCGCTCGCTCGACGCGCGGGCGGGAGGGGCGCGGAGC 158

QY 1231 CCAGCGTCTCCGGCGGTCTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1290
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 159 CCCCCTGGGACCCCGCGGACCCCGCGGACCCCGCGGACCCCGCGGACCCCGCGGACCCCG 218

QY 1291 CGGACCTTAAGACACCTCTGTACCTCCACCCCGCACCC---CACCCACCTCCCGCCAAC 1347
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 219 CGACCCCGGACGCCCGCGGACCCCGCGGACCCCGCGGACCCCGCGGACCCCGCGGACCC 278

QY 1348 TCCCTAGATGTGTGCTGGGCGGTGAAGTCTGCCCGTTTAAGGGGCGGGCGCGCTCCA 1407
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 279 TCCGGGTGCGCCACCTGTGTCTGGCCTCGCGCGCGCGCGCTCGCGCGCGCGCTCGT 338

QY 1408 GTGCTTCTCTGTAGTGAAGTAACT 1433
    | | | | | | | | | | | |
Db 339 GGGCCCGGAGCGGGCGCGGAGCCGGCT 364

RESULT 4
US-09-283-471A-4
: Sequence 4, Application US/09283471A
: Patent No. 6346673
: GENERAL INFORMATION:
: APPLICANT: Roizman, Bernard
: APPLICANT: Chou, Joany
: TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:

```


1	EARLIER APPLICATION NUMBER: 60/038,621	1	EARLIER FILING DATE: 1997-04-11
2	EARLIER FILING DATE: 1997-03-07	2	EARLIER APPLICATION NUMBER: 60/043,677
3	EARLIER APPLICATION NUMBER: 60/040,626	3	EARLIER FILING DATE: 1997-04-11
4	EARLIER FILING DATE: 1997-03-07	4	EARLIER APPLICATION NUMBER: 60/043,315
5	EARLIER APPLICATION NUMBER: 60/040,334	5	EARLIER FILING DATE: 1997-04-11
6	EARLIER FILING DATE: 1997-03-07	6	EARLIER APPLICATION NUMBER: 60/048,974
7	EARLIER APPLICATION NUMBER: 60/040,336	7	EARLIER FILING DATE: 1997-06-06
8	EARLIER FILING DATE: 1997-03-07	8	EARLIER APPLICATION NUMBER: 60/056,886
9	EARLIER APPLICATION NUMBER: 60/040,163	9	EARLIER FILING DATE: 1997-08-22
10	EARLIER FILING DATE: 1997-03-07	10	EARLIER APPLICATION NUMBER: 60/056,877
11	EARLIER APPLICATION NUMBER: 60/047,600	11	EARLIER FILING DATE: 1997-08-22
12	EARLIER FILING DATE: 1997-05-23	12	EARLIER APPLICATION NUMBER: 60/056,889
13	EARLIER APPLICATION NUMBER: 60/047,615	13	EARLIER FILING DATE: 1997-08-22
14	EARLIER FILING DATE: 1997-05-23	14	EARLIER APPLICATION NUMBER: 60/056,893
15	EARLIER APPLICATION NUMBER: 60/047,597	15	EARLIER FILING DATE: 1997-08-22
16	EARLIER FILING DATE: 1997-05-23	16	EARLIER APPLICATION NUMBER: 60/056,630
17	EARLIER APPLICATION NUMBER: 60/047,502	17	EARLIER FILING DATE: 1997-08-22
18	EARLIER FILING DATE: 1997-05-23	18	EARLIER APPLICATION NUMBER: 60/056,878
19	EARLIER APPLICATION NUMBER: 60/047,633	19	EARLIER FILING DATE: 1997-08-22
20	EARLIER FILING DATE: 1997-05-23	20	EARLIER APPLICATION NUMBER: 60/056,662
21	EARLIER APPLICATION NUMBER: 60/047,583	21	EARLIER FILING DATE: 1997-08-22
22	EARLIER FILING DATE: 1997-05-23	22	EARLIER APPLICATION NUMBER: 60/056,872
23	EARLIER APPLICATION NUMBER: 60/047,617	23	EARLIER FILING DATE: 1997-08-22
24	EARLIER FILING DATE: 1997-05-23	24	EARLIER APPLICATION NUMBER: 60/056,882
25	EARLIER APPLICATION NUMBER: 60/047,618	25	EARLIER FILING DATE: 1997-08-22
26	EARLIER FILING DATE: 1997-05-23	26	EARLIER APPLICATION NUMBER: 60/056,637
27	EARLIER APPLICATION NUMBER: 60/047,503	27	EARLIER FILING DATE: 1997-08-22
28	EARLIER FILING DATE: 1997-05-23	28	EARLIER APPLICATION NUMBER: 60/056,903
29	EARLIER APPLICATION NUMBER: 60/047,592	29	EARLIER FILING DATE: 1997-08-22
30	EARLIER FILING DATE: 1997-05-23	30	EARLIER APPLICATION NUMBER: 60/056,888
31	EARLIER APPLICATION NUMBER: 60/047,581	31	EARLIER FILING DATE: 1997-08-22
32	EARLIER FILING DATE: 1997-05-23	32	EARLIER APPLICATION NUMBER: 60/056,879
33	EARLIER APPLICATION NUMBER: 60/047,584	33	EARLIER FILING DATE: 1997-08-22
34	EARLIER FILING DATE: 1997-05-23	34	EARLIER APPLICATION NUMBER: 60/056,880
35	EARLIER APPLICATION NUMBER: 60/047,500	35	EARLIER FILING DATE: 1997-08-22
36	EARLIER FILING DATE: 1997-05-23	36	EARLIER APPLICATION NUMBER: 60/056,894
37	EARLIER APPLICATION NUMBER: 60/047,587	37	EARLIER FILING DATE: 1997-08-22
38	EARLIER FILING DATE: 1997-05-23	38	EARLIER APPLICATION NUMBER: 60/056,911
39	EARLIER APPLICATION NUMBER: 60/047,492	39	EARLIER FILING DATE: 1997-08-22
40	EARLIER FILING DATE: 1997-05-23	40	EARLIER APPLICATION NUMBER: 60/056,636
41	EARLIER APPLICATION NUMBER: 60/047,598	41	EARLIER FILING DATE: 1997-08-22
42	EARLIER FILING DATE: 1997-05-23	42	EARLIER APPLICATION NUMBER: 60/056,874
43	EARLIER APPLICATION NUMBER: 60/047,613	43	EARLIER FILING DATE: 1997-08-22
44	EARLIER FILING DATE: 1997-05-23	44	EARLIER APPLICATION NUMBER: 60/056,910
45	EARLIER APPLICATION NUMBER: 60/047,582	45	EARLIER FILING DATE: 1997-08-22
46	EARLIER FILING DATE: 1997-05-23	46	EARLIER APPLICATION NUMBER: 60/056,864
47	EARLIER APPLICATION NUMBER: 60/047,596	47	EARLIER FILING DATE: 1997-08-22
48	EARLIER FILING DATE: 1997-05-23	48	EARLIER APPLICATION NUMBER: 60/056,631
49	EARLIER APPLICATION NUMBER: 60/047,612	49	EARLIER FILING DATE: 1997-08-22
50	EARLIER FILING DATE: 1997-05-23	50	EARLIER APPLICATION NUMBER: 60/056,845
51	EARLIER APPLICATION NUMBER: 60/047,632	51	EARLIER FILING DATE: 1997-08-22
52	EARLIER FILING DATE: 1997-05-23	52	EARLIER APPLICATION NUMBER: 60/056,892
53	EARLIER APPLICATION NUMBER: 60/047,601	53	EARLIER FILING DATE: 1997-08-22
54	EARLIER FILING DATE: 1997-05-23	54	EARLIER APPLICATION NUMBER: 60/057,761
55	EARLIER APPLICATION NUMBER: 60/043,580	55	EARLIER FILING DATE: 1997-08-22
56	EARLIER FILING DATE: 1997-04-11	56	EARLIER APPLICATION NUMBER: 60/047,595
57	EARLIER APPLICATION NUMBER: 60/043,568	57	EARLIER FILING DATE: 1997-05-23
58	EARLIER FILING DATE: 1997-04-11	58	EARLIER APPLICATION NUMBER: 60/047,599
59	EARLIER APPLICATION NUMBER: 60/043,314	59	EARLIER FILING DATE: 1997-05-23
60	EARLIER FILING DATE: 1997-04-11	60	EARLIER APPLICATION NUMBER: 60/047,588
61	EARLIER APPLICATION NUMBER: 60/043,569	61	EARLIER FILING DATE: 1997-05-23
62	EARLIER FILING DATE: 1997-04-11	62	EARLIER APPLICATION NUMBER: 60/047,585
63	EARLIER APPLICATION NUMBER: 60/043,311	63	EARLIER FILING DATE: 1997-05-23
64	EARLIER FILING DATE: 1997-04-11	64	EARLIER APPLICATION NUMBER: 60/047,586
65	EAR		


```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163...4099
; US-09-041-886-18

Query Match 2.5%; Score 38.6; DB 4; Length 4481;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 1187 CTGCGGGAGCTGACTAGAGAGTCTCGCGCGAGCCCGGAGCCCGCGCTTCCCGGCG 1246
Db 194 CCGGGTGGCCACCGAGCTCGCGCGCTTCGCGCGAGCCGAGTGGCGGGTGGCGCTCGC 253
Qy 1247 GCTTAGCGCGCGCGCGCGCGGAGGAGGAGGAGCGGAGCGGAGCCCTAAGACACC 1306
Db 254 TCCAGCGCGCGCGCGCGGAGCGGCGGCGGCGGCTGGCGCGCGCGCGGACCGGTATC 313
Qy 1307 TCGGTACCTTCCACCCCGGAGCCCGGAGCCCGGAGCCCGGAGCCCGGAGCCCGG 1351
Db 314 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 358

RESULT 11
US-08-709-838-1
; Sequence 1, Application US/08/709838
; Patent No. 6140064
; GENERAL INFORMATION:
; APPLICANT: Loetscher, Marcel
; APPLICANT: Moser, Bernhard
; TITLE OF INVENTION: IP-10/MIG RECEPTOR DESIGNATED CXCR3,
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USES THEREFOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,838
; FILING DATE: 31-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/709,838
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook Esq., David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: TKI96-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1670 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69...1172
; US-08-829-839-1

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1670 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69...1172
; US-08-709-838-1

Query Match 2.5%; Score 38.2; DB 3; Length 1670;
Best Local Similarity 59.8%; Pred. No. 0.78;
Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 714 TCCTGCTTTTATCAGAGGAGCTGATCAATATATGAATTAAGAGGGGGTGGTCC 773
Db 1562 TCTTTTATTTTATGTAAATCCIGCTTAAACTTTCATTAACAAGATCGTCAGGA 1621
Qy 774 CATATTGTCGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 820
Db 1622 CTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1668
```

```
Query Match          2.5%; Score 38.2; DB 4; Length 1670;
Best Local Similarity 59.8%; Pred No. 0.78; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 43;

QY 714 TCTGTGTTTATCATCAGGAGGCTGATCAATATATGAAATTAAGGGGCTGCTCC 773
      ||| ||||| | | ||| | | | ||| | | | ||| | | | ||| |
Db 1562 TCTTTTATTATGCTAAATCCGCTAAACCTTTTCAATAAACAAGATGTCAGGA 1621

QY 774 CATATGCTCTGCTGTTTGTGTTGTTGTTCTCTCTCTCTCTCTCTCTCTCT 820
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1622 CCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1668

RESULT 13
US-09-345-882-1/C
; Sequence 1, Application US/09345882
; Patent No. 6395373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSEI 1031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,509
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103606
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
```

```
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134134
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134374
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72711..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72711..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
```


FEATURE:
NAME/KEY: allele
LOCATION: 9390..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108054..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108054..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match 2.5%; Score 37.6; DB 4; Length 162450;
Best Local Similarity 52.2%; Pred. No. B.4;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 686 TCACAAATGATACAAACTAAATACAAAGTCTGTGTATATATACAGGAGGCGATCAA 745
DB 109593 TAAACATTTTTTACTATATATATATAGCCCAATTTTATTTCCCTCCCTGTTCTG 109534
QY 746 TATAATGAATATAAAGGGGCTGCTCCATATATGTTCTGTGTTTGTGTTTGTTC 805
DB 109533 CTGCATACTGIGGCCAAATGTCAGTACGCTAAATGATTTTTTTTTTTTTTTT 109474
QY 806 TTTTGTGTTTGTGCTCCCTCTCTCAATTAAGAAG 846
DB 109473 TTTTGTGTTTGTGCTCCCTCTCTCAATTAAGAAG 109433
RESULT 14
US-08-018-977C-4/c
; Sequence 4, Application JS/08018977C
; Patent No. 5686601
; GENERAL INFORMATION:
; APPLICANT: Weber, Peter C.
; TITLE OF INVENTION: DNA Sequences Encoding Mutant Antiviral
; TITLE OF INVENTION: Regulatory Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: The Pennsylvania State University
; STREET: 113 Technology Center
; CITY: University Park
; STATE: PA
; COUNTRY: US
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS 6.22
; SOFTWARE: Microsoft Word for Windows 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/018,977C
; FILING DATE: 18-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/726,071
; FILING DATE: 05-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas J. Monahan
; REGISTRATION NUMBER: 29,835
; REFERENCE/DOCKET NUMBER: 9-1039/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (814) 865-6277
; TELEFAX: (814) 865-3591
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 bases
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-018-977C-4

Query Match 2.5%; Score 37.6; DB 1; Length 936;
Best Local Similarity 54.3%; Pred. No. 0.87;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 1212 CTGGGGCGCAGCCCGGAGCGCTTCCCGCGGCTTAGAGCGGGCGCGCGGGGG 1271
DB 785 CTCGAGCGACCGCGCCCTCTCTGCTGACAGAGGCGGGGAGGCGGGCGCGCG 726
QY 1272 GGAAGGGGAGCAGACCGCGGACCTTACAGACACTGTACCTCCACCCACCCCA 1331
DB 725 GAGGGGGGCGCGCGGGAGGGCGCGTCCACCTCCACCGCGCGCGCGCGCGCG 666
QY 1332 CCCAGCTCCCGCAACTCC 1351
DB 665 CGCGCGCGCTCGCGCGCGCG 646

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:46:48 ; Search time 154.036 Seconds
(without alignments)
8245.650 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1080_1643

Perfect score: 564

Sequence: 1 gagctctctctcccccaatc.....caggggcaggggggagctc 564

Scoring table: IDENTITY-NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*

```

1: /SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	564	100.0	1643	22	AAF24681
2	564	100.0	1643	22	AAF24703
3	552	97.9	1197	24	ABL58400
4	548	97.2	183999	22	AAF92831
5	547.4	97.1	3231	24	AAF37265
6	473.4	83.9	2910	24	AAF37267
7	119.4	21.2	227	21	AA09615
8	98	17.4	7260	22	AA021326
9	98	17.4	7260	22	AA170315

10	91	16.1	221	24	AAF24681
11	89.4	15.9	736	22	AAF24703
12	89.4	15.9	1556	22	AAF18606
13	89	15.8	1750	22	AAH04729
14	89	15.8	1750	22	AAH17451
15	77.2	13.7	7281	22	AAK51683
16	75.4	13.4	7086	22	ABA09200
17	75.4	13.4	7086	22	AAK52667
18	75	13.3	9854	22	AAK52667
19	67	11.9	10442	22	AAK52667
20	67	11.9	10442	22	AAK52667
21	60	10.6	10474	22	AAK52667
22	60	10.6	10474	22	AAK52667
23	60	10.6	10474	22	AAK52667
24	60	10.6	10474	22	AAK52667
25	45	8.0	114955	20	AAK53491
26	38.8	6.9	595	14	AAQ50041
27	38.8	6.9	1335	13	AAQ50041
28	38.6	6.8	4367	19	AAV3027C
29	38.6	6.8	4481	19	AAV06552
30	38.6	6.8	4481	20	AAZ23428
31	38.6	6.8	18660	21	AAA58472
32	37.6	6.7	381	24	ABN73567
33	37.6	6.7	936	18	AAQ97101
34	37.4	6.6	1599	16	AAQ97101
35	37.4	6.6	12042	23	AAK76999
36	37.4	6.6	175737	24	ABK83571
37	37	6.6	1555	21	AAA64196
38	36.8	6.5	565	8	AAW70930
39	36.8	6.5	565	22	AAF32043
40	36.8	6.5	19408	22	AAK42003
41	36.8	6.5	19408	22	AAK87230
42	36.8	6.5	19408	22	AAK90644
43	36.6	6.5	128139	21	AAK64291
44	36.6	6.5	133719	21	AAK64754
45	36.4	6.5	38	22	AAF93064

ALIGNMENTS

RESULT: 1
AAF24681
ID AAF24681 standard; DNA; 1643 bp.
XX
AC AAF24681;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
PN WC200078972-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16765.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
XX
PI Lawn RM, Wade D, Garvin M;
XX
DR WPI; 2001-37812/14.

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
PI useful for the development of agents for the treatment of heart disease
PI and other disorders associated with hypercholesterolemia and
PT atherosclerosis -
XX
PS Claim 1: Page 143-144; 215pp; English.
XX
CC The present sequence represents the 5' flanking region of the human
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
CC resides in cell membranes and utilises ATP hydrolysis to transport a wide
CC variety of substrates across the plasma membrane. ABC1 is a pivotal
CC protein in the apolipoprotein-mediated mobilisation of intracellular
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
CC are useful for developing pharmaceutical agents for the treatment of
CC heart disease and other disorders associated with hypercholesterolemia
CC and atherosclerosis. The genes are useful for developing screening assays
CC to screen for compounds that regulate the expression of genes associated
CC with cholesterol transport. The genes and proteins are also useful for
CC other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 100.0%; Score 564; DB 22; Length 1643;
Best Local Similarity 100.0%; Pred. No. 3.9e-140;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAT 60
DB 1080 GAGCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAT 1139
QY 61 TCGGGAAAGCAGGATTTAGAGGAAGAAATTCACACTGGTCCCTTGGCTCCCGGAACT 120
DB 1140 TCGGGAAAGCAGGATTTAGAGGAAGAAATTCACACTGGTCCCTTGGCTCCCGGAACT 1199
QY 121 GGACTAGAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTCCGCGCGCTTAGCGCGG 180
DB 1200 GGACTAGAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTCCGCGCGCTTAGCGCGG 1259
QY 181 GGGCCCGGGGGGGAAGGGGAGCGAGACCGGACCCCTAAGACACCTGTGTACCCCTCC 240
DB 1260 GGGCCCGGGGGGGAAGGGGAGCGAGACCGGACCCCTAAGACACCTGTGTACCCCTCC 1319
QY 241 ACCCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 300
DB 1320 ACCCCACCCACCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1379
QY 301 CCGGTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 360
DB 1380 CCGGTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1439
QY 361 ACAGAGCCCGGGAACGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 1440 ACAGAGCCCGGGAACGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1499
QY 421 CTGGCTCTGGTGGAGCGGGAATCTATAAAGAACTATCTCCCGGCAAAACCCGTAATG 480
DB 1500 CTGGCTCTGGTGGAGCGGGAATCTATAAAGAACTATCTCCCGGCAAAACCCGTAATG 1559
QY 481 CGAGCGAGAGTGTAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 540
DB 1560 CGAGCGAGAGTGTAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 619
QY 541 GCGGCAAGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 564
DB 1620 GCGGCAAGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1643

RESULT 2
AAF24703

ID
XX AAF24703 standard; DNA; 1643 BP.
AC AAF24703;
XX
XX 20-APR-2001 (first entry)
DT
XX
DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
KW Human: adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
PN WO200078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16591.
XX
XX 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
XX (ONIW) UNIV WASHINGTON.
XX
XX Lawn RM, Wade D, Oram JF, Garvin M;
XX WPI; 2001-137811/14.
XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX polynucleotides and polypeptides, useful for treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX
XX Disclosure; Page 138-139; 211pp; English.
XX
XX The present sequence represents the 5' flanking region of the human
XX adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
XX resides in cell membranes and utilises ATP hydrolysis to transport a wide
XX variety of substrates across the plasma membrane. ABC1 is a pivotal
XX protein in the apolipoprotein-mediated mobilisation of intracellular
XX cholesterol stores. ABC1 is defective in Tangier disease, a genetic
XX disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
XX gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
XX are useful for developing pharmaceutical agents for the treatment of
XX heart disease and other disorders associated with hypercholesterolemia
XX and atherosclerosis. The genes are useful for developing screening assays
XX to screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX other disorders associated with hypercholesterolemia.
XX
SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 100.0%; Score 564; DB 22; Length 1643;
Best Local Similarity 100.0%; Pred. No. 3.9e-140;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAT 60
DB 1080 GAGCTCTCTCTCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAT 1139
QY 61 TCGGGAAAGCAGGATTTAGAGGAAGAAATTCACACTGGTCCCTTGGCTCCCGGAACT 120
DB 1140 TCGGGAAAGCAGGATTTAGAGGAAGAAATTCACACTGGTCCCTTGGCTCCCGGAACT 1199
QY 121 GGACTAGAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTCCGCGCGCTTAGCGCGG 180
DB 1200 GGACTAGAGAGTCTGGCGGCGAGCCCGAGCCGAGCGCTTCCGCGCGCTTAGCGCGG 1259

QY 181 GGGCCGGGGGGGAGGGGAGAGCCCGGAGCCCTAAGACACCTCTGTACCTCC 240
DB 1260 GGGCCGGGGGGGAGGGGAGAGCCCGGAGCCCTAAGACACCTCTGTACCTCC 1319
QY 241 ACCCCACCCACCCACCTCCCGGAGAGAGCCCTAAGACACCTCTGTACCTCC 300
DB 1320 ACCCCACCCACCCACCTCCCGGAGAGAGCCCTAAGACACCTCTGTACCTCC 1379
QY 301 CCGGTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 360
DB 1380 CCGGTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1439
QY 361 ACAGAGCCGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420
DB 1440 ACAGAGCCGGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1499
QY 421 CTGGGTCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 1500 CTGGGTCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559
QY 481 CGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 1560 CGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1619
QY 541 CGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 564
DB 1620 CGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1643

RESULT 3
ABL58400
ID ABL58400 standard; DNA; 1197 BP.

AC ABL58400:
XX
XX
XX 30-JUL-2002 (first entry)
DE Human large ATP-binding cassette transporter 1(hABCl), promoter sequence.
XX Human: large ATP-binding cassette transporter 1; ABCL; promoter;
XX antiarteriosclerotic; gene transfer; transactivator; ds.
XX Homo sapiens.
XX WO200183506-A1.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13654.
XX 28-APR-2000; 2000US-0560372.
XX (UYCO) UNIV COLUMBIA NEW YORK.
XX Tall AR;
XX WPI: 2002-049334/06.
XX Novel isolated human large ATP-binding cassette transporter 1 promoter
XX capable of directing transcription of heterologous coding sequence
XX positioned downstream to it, useful for expressing foreign DNA in host
XX cells
XX
XX Claim 1; Fig 3; 68pp; English.
XX
XX The invention relates to an isolated human large ATP-binding cassette
XX transporter 1 (ABCL) promoter capable of directing transcription of
XX heterologous coding sequence positioned downstream to it. The hABCL
XX promoter is useful for expressing foreign DNA in a host cell, by
XX introducing into the host cell a gene transfer vector comprising the
XX promoter operably linked to a foreign DNA encoding a desired polypeptide
XX or RNA, where the foreign DNA is expressed. The gene transfer can be
XX introduced into the host cell by adenovirus infection, liposome-mediated

CC transfer, topical application to the cell or microinjection. The gene
CC transfer vector encodes and expresses a reporter molecule. The method
CC further involves introducing into the cell a gene transfer vector
CC comprising a nucleic acid segment encoding a transactivator protein
CC capable of upregulating the ABCL promoter, or contacting the cell with
CC the transactivator protein, or an agonist of the transactivator protein.
CC Modulators of human ABCL gene expression are useful for treating
CC atherosclerosis. The present sequence represents the hABCL promoter.
XX
XX

Sequence 1197 BP: 284 A; 314 C; 328 G; 271 T; 0 other;

Query Match 97.9%; Score 552; DB 24; Length 1197;
Best Local Similarity 100.0%; Pred. No. 5,6e-137;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAATTCGGAAGCAG 72
DB 630 CCCCAATCCCTCCCTCCGGCTGAGGAACTAACAAAGGAAAAAAATTCGGAAGCAG 689
QY 73 GATTAGAGGAAGCAATTCACCTGGTGGCTGGGAGGAACTGACTAGACAGT 132
DB 690 GATTAGAGGAAGCAATTCACCTGGTGGCTGGGAGGAACTGACTAGACAGT 749
QY 133 CTGCGGCGCAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 192
DB 750 CTGCGGCGCAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
QY 193 GGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
DB 810 GGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
QY 253 CCCACCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
DB 870 CCCACCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
QY 313 GCGGGCGGCGGCTCCAGCTGCTTTCTGCTGAGTGAAGTGAAGTGAAGTGAAG 372
DB 930 GCGGGCGGCGGCTCCAGCTGCTTTCTGCTGAGTGAAGTGAAGTGAAGTGAAG 989
QY 373 AACGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
DB 990 AACGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049
QY 433 CAGCGGAATCTATAAAGGAACTAGTCCCGGAGAGAGAGAGAGAGAGAGAGAG 492
DB 1050 CAGCGGAATCTATAAAGGAACTAGTCCCGGAGAGAGAGAGAGAGAGAGAGAG 1109
QY 493 AGTGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
DB 1110 AGTGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1169
QY 553 GCGGGGAGAGCTC 564
DB 1170 GCGGGGAGAGCTC 1181

RESULT 4
AAF92831
ID AAF92831 standard; DNA; 183999 BP.
XX
XX AAF92831;
XX
XX 17-MAY-2001 (first entry)
XX Human ABCL genomic DNA.
XX
XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABCL; ds.
XX Homo sapiens.
XX WO200115676-A2.
XX
XX 08-MAR-2001.

XX PF 01-SEP-2000; 2000WO-IB01492.
XX XX
XX 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WPI: 2001-244356/25.
XX
XX Treating a lower than normal high density lipoprotein-cholesterol
PT (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
XX Claim 8; Fig 1; 317pp; English.
XX
XX The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal
CC triglyceride level, or a cardiovascular disease, involving
CC administering a compound that modulates LXR- or RXR-mediated
CC transcriptional activity or ABC1 expression or activity.
CC The LXR gene product may be used in an assay to identify
CC compounds useful for the treatment of a disease or condition selected a
CC lower than normal HDL cholesterol level, a higher than normal
CC triglyceride level, and a cardiovascular disease.
XX
XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;
SQ

Query Match 97.2%; Score 548; DB 22; Length 183999;
Best Local Similarity 98.9%; Pred. No. 2.8e-135;
Matches 563; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
XX 1 GAGCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACATAACAAAGGAAAAAATAAT 60
Db 28312 GAGCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACATAACAAAGGAAAAAATAAT 28371
XX 51 TCGGAAACAGAGATTAGAGGAAGCAATTCACCTGGTCCCTTGGCTGCGGGAACGT 120
Db 26372 TCGGAAACAGAGATTAGAGGAAGCAATTCACCTGGTCCCTTGGCTGCGGGAACGT 28431
XX 121 GGACTAGAGAGTCTCGGGGACGCCCGAGCCACGCTTCCCGCGGCTCTTAGCCCGGC 180
Db 28432 GGACTAGAGAGTCTCGGGGACGCCCGAGCCACGCTTCCCGCGGCTCTTAGCCCGGC 28491
XX 181 GGGCCCGGGGGGAAAGGGGACGAGACCGGGACCTTAAGACACCTGCTACCTCC 240
Db 28492 GGGCCCGGGGGGAAAGGGGACGAGACCGGGACCTTAAGACACCTGCTACCTCC 28551
XX 24 A-----CCCCACCCACCACTCCGCCCACTCCCTAGATGTGCTGGCGGCTGAA 295
Db 28552 ANNNNNCCCCACCCACCACTCCGCCCACTCCCTAGATGTGCTGGCGGCTGAA 28611
XX 296 CGTCCGCGCTTTAAGGGGCGCGCGGCTCCACGTCCTTCTGCTGAGTGAAGTA 355
Db 28612 CGTCCGCGCTTTAAGGGGCGCGCGGCTCCACGTCCTTCTGCTGAGTGAAGTA 28671
XX 356 CATAACAGAGCGCGGAACGGGGGCGGGAGGAGGAGACAGGCTTTGACCGTAGT 415
Db 28672 CATAACAGAGCGCGGARACGGGGGCGGGAGGAGGAGGAGGAGGAGGAGGAGT 28731
XX 416 AAGCTCTGGCTGGTTCAGCCGAATCTATAAAGGAATAGTCCCGGCAAAACCCCGT 475
Db 28732 AAGCTCTGGCTGGTTCAGCCGAATCTATAAAGGAATAGTCCCGGCAAAACCCCGT 28791
XX 476 AATTGGCAGGACAGTGTGGGGCGGGACCCCGAGACCGGACCGCTCTCTCC 535
Db 28792 AATTGGCAGGACAGTGTGGGGCGGGACCCCGAGACCGGACCGGACCGCTCTCTCC 28851

QY 536 GGCTGGCGCAGGCGAGGCGGGGAGC 564
Db 28852 GGCTGGCGCAGGCGAGGCGGGGAGCTC 28880
RESULT 5
AAD37265
ID AAD37265 standard; DNA; 3231 3P.
XX AC
XX AAD37265;
XX DT 21-AUG-2002 (first entry)
XX DE Human ABC1 transcription regulatory DNA #1.
XX KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
XX KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; GS.
XX OS Homo sapiens.
XX PN W0200183746-A2.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-EP05488.
XX PR 02-MAY-2000; 2000US-201280P.
XX PA (AVET) AVENTIS PHARMA SA.
XX PI Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
XX PI Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
XX WPI: 2002-154404/20.
XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
XX (ABC1) and screening for candidate modulatory compounds or substances
XX
XX Claim 1; Page 130-131; 152pp; English.
XX The invention relates a nucleic acid which is capable of regulating the
XX transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
XX casual gene for pathologies linked to a dysfunctioning of cholesterol
XX metabolism, including diseases such as atherosclerosis. Polynucleotides
XX of the invention are used to screen candidate molecules or substances
XX that are capable of modulating the transcription of the ABC1 gene. They
XX are used in antisense therapy. Compositions comprising sequences of the
XX invention are used to treat hypercholesterolaemia and atherosclerosis.
XX The present sequence is human ABC1 transcription regulating DNA.
SQ Sequence 3231 BP; 809 A; 773 C; 876 G; 773 T; 0 other;

Query Match 97.1%; Score 547.4; DB 24; Length 3231;
Best Local Similarity 98.9%; Pred. No. 1.2e-135;
Matches 563; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 GAGCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACATAACAAAGGAAAAAATAAT 60
Db 2415 GAGCTCTCTCTCCCAATCCCTCCCTCCGCTGAGGAACATAACAAAGGAAAAAATAAT 2475
QY 61 TCGGAAACAGAGATTAGAGGAAGCAATTCACCTGGTGGCTTGGCTGCGGGAACGT 120
Db 2476 TCGGAAACAGAGATTAGAGGAAGCAATTCACCTGGTGGCTTGGCTGCGGGAACGT 2535
QY 121 GGACTAGAGAGTCTCGGGGCGGACCCCGAGCCGCTTCCCGCGGCTTAGCCCGGC 180
Db 2536 GGACTAGAGAGTCTCGGGGCGGACCCCGAGCCGCTTCCCGCGGCTTAGCCCGGC 2595
QY 181 GGGCCCGGGGGGAAAGGGGACGAGACCGGGACCTTAAGACACCTGCTGTACCTCC 240
Db 2596 GGGCCCGGGGGGAAAGGGGACGAGACCGGGACCTTAAGACACCTGCTGTACCTCC 2655

```
OY 241 AC-----CCCCACCCACCCACCTCCCCCACCTCCCTAGATGTCGCGGGGGCGTGAA 295
D 2656 ACCCCACCCACCCACCCACCCACCTCCCCCACCTCCCTAGATGTCGCGGGGGCGTGAA 2715
OY 296 CGTCGCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 355
D 2716 CGTCGCCGCTTTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2775
OY 356 CATAAAGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 415
D 2776 CATAAAGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2835
OY 416 AACCTCTGCGCTCGGTCACGCGCAATCTATAAAGGAAGTATCCCGGCAAAACCCCGT 475
D 2836 AACCTCTGCGCTCGGTCACGCGCAATCTATAAAGGAAGTATCCCGGCAAAACCCCGT 2895
OY 476 AATTGCGAGGAGAGTGTAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 535
D 2896 AATTGCGAGGAGAGTGTAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2955
OY 536 GGGTCGCGGAGGCGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 564
D 2956 GGGTCGCGGAGGCGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2984

RESULT 6
AAD37267
ID AAD37267 standard; DNA; 2910 bp.
XX
AC AAD37267;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #3.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT protein_bind 12..23
FT /*tag= a
FT /bound_moiety= "LMO2COM/MYOD"
FT
FT protein_bind 97..107
FT /*tag= b
FT /bound_moiety= "DeltaE1"
FT
FT protein_bind 110..125
FT /*tag= c
FT /bound_moiety= "S8/NKX2.5"
FT
FT protein_bind 196..211
FT /*tag= d
FT /bound_moiety= "S8"
FT
FT protein_bind 228..237
FT /*tag= e
FT /bound_moiety= "GATA"
FT
FT protein_bind 399..410
FT /*tag= f
FT /bound_moiety= "IK2"
FT
FT protein_bind 412..420
FT /*tag= g
FT /bound_moiety= "LYF1"
FT
FT protein_bind 528..539
FT /*tag= h
FT /bound_moiety= "LMO2COM/MYOD/DeltaE1"
FT
FT protein_bind 549..556
FT /*tag= i
FT /bound_moiety= "LYF1"
FT
FT protein_bind 558..568
FT /*tag= j
FT /bound_moiety= "DeltaE1"
FT
FT protein_bind 590..596
```

```
FT /*tag= k
FT /bound_moiety= "LXX2.5"
FT
FT protein_bind 688..620
FT /*tag= l
FT /bound_moiety= "NFY/CAAT"
FT
FT protein_bind 708..715
FT /*tag= m
FT /bound_moiety= "MZF1"
FT
FT protein_bind 723..730
FT /*tag= n
FT /bound_moiety= "MZF1"
FT
FT protein_bind 771..785
FT /*tag= o
FT /bound_moiety= "HPH2/SRY/ENV1"
FT
FT protein_bind 803..812
FT /*tag= p
FT /bound_moiety= "CREB3L1/VBP"
FT
FT protein_bind 831..837
FT /*tag= r
FT /bound_moiety= "NKX2.5"
FT
FT protein_bind 1076..1089
FT /*tag= s
FT /bound_moiety= "GATA"
FT
FT protein_bind 1173..1188
FT /*tag= t
FT /bound_moiety= "LXRalpha/deltaE1"
FT
FT protein_bind 1189..1209
FT /*tag= u
FT /bound_moiety= "DeltaE1/LYF1/IK2"
FT
FT protein_bind 1483..1491
FT /*tag= v
FT /bound_moiety= "AP4"
FT
FT protein_bind 1498..1514
FT /*tag= w
FT /bound_moiety= "LMO2-COM/MYOD/deltaE1/B47"
FT
FT protein_bind 1524..1545
FT /*tag= x
FT /bound_moiety= "ZID/deltaE1"
FT
FT protein_bind 1597..1607
FT /*tag= y
FT /bound_moiety= "DeltaE1"
FT
FT protein_bind 1622..1627
FT /*tag= z
FT /bound_moiety= "PPAR"
FT
FT protein_bind 1632..1637
FT /*tag= aa
FT /bound_moiety= "PPAR"
FT
FT protein_bind 1685..1698
FT /*tag= ab
FT /bound_moiety= "USF/NMYC/MYCNAX"
FT
FT protein_bind 1787..1797
FT /*tag= ac
FT /bound_moiety= "DeltaE1"
FT
FT protein_bind 1809..1819
FT /*tag= ad
FT /bound_moiety= "DeltaE1"
FT
FT protein_bind 1822..1833
FT /*tag= ae
FT /bound_moiety= "SRY"
FT
FT protein_bind 1840..1850
FT /*tag= af
FT /bound_moiety= "AP1"
FT
FT protein_bind 1942..1956
FT /*tag= ag
FT /bound_moiety= "HNF3beta"
FT
FT protein_bind 1978..1985
FT /*tag= ah
FT /bound_moiety= "NKX2.5"
FT
FT protein_bind 2008..2016
FT /*tag= ai
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT
FT protein_bind 2019..2024
FT /*tag= aj
```

```

FT protein_bind /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2051..2059 /*tag= ak
FT /bound_moiety= "GATA"
FT 2104..2111 /*tag= al
FT /bound_moiety= "SOX5"
FT 2114..2152 /*tag= am
FT /bound_moiety= "SYR/HFH/HNF3beta"
FT 2221..2228 /*tag= an
FT /bound_moiety= "MZF1"
FT 2236..2248 /*tag= ao
FT /bound_moiety= "IK2/NFKappaB/CREL"
FT 2259..2272 /*tag= ap
FT /bound_moiety= "LMO2COM/GATA"
FT 2289..2306 /*tag= aq
FT /bound_moiety= "MZF1/SRY"
FT 2313..2318 /*tag= ar
FT /bound_moiety= "PPAR"
FT 2321..2326 /*tag= as
FT /bound_moiety= "PPAR"
FT 2335..2342 /*tag= at
FT /bound_moiety= "MZF1"
FT 2361..2384 /*tag= av
FT /bound_moiety= "MZF1"
FT 2442..2451 /*tag= au
FT /bound_moiety= "HNF3beta/SRY/EVIL"
FT 2426..2433 /*tag= ax
FT /bound_moiety= "MZF1"
FT 2455..2466 /*tag= ay
FT /bound_moiety= "MZF1"
FT 2491..2498 /*tag= az
FT /bound_moiety= "STAT"
FT 2524..2534 /*tag= ba
FT /bound_moiety= "PPAR"
FT 2536..2541 /*tag= bb
FT /bound_moiety= "PPAR"
FT 2589..2600 /*tag= bc
FT /bound_moiety= "PPAR"
FT 2610..2617 /*tag= bd
FT /bound_moiety= "PPAR"
FT 2634..2648 /*tag= be
FT /bound_moiety= "PPAR"
FT 2657..2672 /*tag= bf
FT /bound_moiety= "PPAR"
FT 2728..2740 /*tag= bg
FT /bound_moiety= "PPAR"
FT 2743..2757 /*tag= bh
FT /bound_moiety= "PPAR"

```

```

FT protein_bind 2758..2773 /*tag= bi
FT /bound_moiety= "NFE2A1"
FT 2774..2787 /*tag= bj
FT /bound_moiety= "XFD1/HFH"
FT 2794..2806 /*tag= bk
FT /bound_moiety= "GC/SP1/MZF1"

Query Match 83.9%; Score 473.4; DP 24; Length 2910;
Best Local Similarity 96.8%; Pred. No. 5.6e-116;
Matches 489; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 GAGCTCTCTC-CCCCCAATCCCTCCGCTGAGGAACTAACAAAGGAAAAAAT 60
Db 2416 GAGCTCTCTCTCCCCCAATCCCTCCGCTGAGGAACTAACAAAGGAAAAAAT 2475
QY 61 TCCGGAAGACAGGATTAGAGGAAGCAAAATTCACACTGGTGCCTTGGCTGCCGGGAAGT 120
Db 2476 TCCGGAAGACAGGATTAGAGGAAGCAAAATTCACACTGGTGCCTTGGCTGCCGGGAAGT 2535
QY 121 GGACTAGAGAGTCTCGGGCGAGCCCGGAGCCCGGCTTCCCGCGGCTCTTAGCCCGC 180
Db 2536 GGACTAGAGAGTCTCGGGCGAGCCCGGAGCCCGGCTTCCCGCGGCTCTTAGCCCGC 2595
QY 181 GGGCCCGGGGGGGAAGGGGACGACGCGGAGCCCTAAGACACCTGCCTGACCCCTCC 240
Db 2596 GGGCCCGGGGGGGAAGGGGACGACGCGGAGCCCTAAGACACCTGCCTGACCCCTCC 2655
QY 241 AC-----CCCCACCCACCCACCTCCCGCCAACTCCCTAGATGTCTGTGGGCGCTGAA 295
Db 2656 ACCCCACCCACCCACCCACCCACCTCCCGCCAACTCCCTAGATGTCTGTGGGCGCTGAA 2715
QY 296 CGTCGCCGCTTAAAGGGGGGGGGCGGCTCCACGCTCTTCTGCTGAGTACTGAACCTA 355
Db 2716 CGTCGCCGCTTAAAGGGGGGGGGCGGCTCCACGCTCTTCTGCTGAGTACTGAACCTA 2775
QY 356 CATAAACAGAGCCGGGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 415
Db 2776 CATAAACAGAGCCGGGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 2835
QY 416 AACCTCTGCGCTCGGTGACGCGCAATCTATAAAGGAAGTACTCCCGGCAAAACCCCGT 475
Db 2836 AACCTCTGCGCTCGGTGACGCGCAATCTATAAAGGAAGTACTCCCGGCAAAACCCCGT 2895
QY 476 AATTGGAGCGGAGAG 490
Db 2896 AATTGGAGCGGAGAG 2910

RESULT 7
AAC09615
ID AAC09615 standard; cDNA; 227 BP.
XX AAC09615;
XX AAC09615;
XX 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 13690.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 990S-0122487.

```


PT variation /*tag= e
 FT replace(3836,C)
 FT /*tag= f

XX EP1136554-A1.

XX 26-SEP-2001.

PF 24-MAR-2000; 2000EP-C106401.

XX 24-MAR-2000; 2000EP-0106401.

XX (FARB) BAYER AG.

XX Schmitz G, Bodzioch M;

XX WPI; 2001-640389/74.

DR P-ESDB; AAM50228.

XX New adenosine triphosphate binding cassette transporter gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 FT cardiovascular diseases and inflammatory diseases

XX Disclosure: Page 26-28; 41pp; English.

XX The present sequence is that of cDNA encoding the human adenosine
 CC triphosphate (ATP) binding cassette transporter 1 (ABCI) protein
 CC (see AAM50228). The sequence includes an extended open reading
 CC frame (ORF) to that provided by the sequence in AAI70314, using
 CC an alternative ATG codon as initiation codon and thereby adding an
 CC extra 40 N-terminal amino acids to the encoded ABCI protein (see
 CC AAM50228). The invention provides 4 common polymorphisms in the
 CC ABCI gene. These were identified by sequencing the ABCI gene in
 CC different Tangier kindreds. In the variant genes (numbering as in
 CC AAI70314), G is changed to A at position 536, T is changed to C at
 CC position 1136, A is changed to G at position 2589 or G is changed
 CC to C at position 3456, or any combination of these. All of these
 CC polymorphisms alter the amino acid sequence of ABCI and therefore
 CC may affect its function. The 2 most common polymorphisms (G596A)
 CC and A2589G are both associated with a decreased in vitro ApoA-I
 CC mediated efflux of cholesterol from mononuclear phagocytes, a
 CC feature typical of Tangier disease. 3 Of the variants (G596A,
 CC A2589G and G3456C) are significantly increased in a population of
 CC men having low high density lipoprotein-cholesterol levels and
 CC established coronary heart disease (CHD) relative to CHD-free
 CC control subjects. The use of the provided ABCI polymorphisms for
 CC the diagnosis and treatment of lipid disorders, cardiovascular
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus
 CC erythematoses) is claimed. Modulation of ABCI transcripts or
 CC proteins by antisense or ribozyme technology or RNA decoys is also
 CC claimed.

SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match: 17.4%; Score 98; DB 22; Length 7260;

Best Local Similarity 100.0%; Pred. No. 4.4e-16;

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 AAACCCGCGTAATTGCGAGCGAGAGTACGTGGCGCGGGACCGCGAGCGCGACCC 526

DB 1 AAACCCGCGTAATTGCGAGCGAGAGTACGTGGCGCGGGACCGCGAGCGCGACCC 50

QY 527 TTCTCTCCCGGGCTGGCGAGCGAGCGGGGGAGCTC 564

DB 61 TTCTCTCCCGGGCTGGCGAGCGAGCGGGGGAGCTC 98

RESULT 10

AAD37268

ID AAD37268 standard; DNA; 221 BP.

XX

AC AAD37268;

XX

21-AUG-2002 (first entry)

XX Human ABCI gene exon 1A.

XX Human; ATP-binding cassette 1; ABCI gene regulation; atherosclerosis;
 KW cholesterol metabolism; hypercholesterolaemia; antisense therapy;
 KW exon 1A; ds.

XX Homo sapiens.

XX WO200183746-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-EP05488.

XX 02-MAY-2000; 2000US-201280P.

XX (AVET) AVENTIS PHARMA SA.

XX Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;

XX Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;

XX WP; 2002-154404/20.

XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
 PT (ABCI) and screening for candidate modulatory compounds or substances

XX Claim 4; Page 132; 152pp; English.

XX The invention relates a nucleic acid which is capable of regulating the
 CC transcription of human ATP-binding cassette 1 (ABCI) gene, which is a
 CC causal gene for pathologies linked to a dysfunctioning of cholesterol
 CC metabolism, including diseases such as atherosclerosis. Polynucleotides
 CC of the invention are used to screen candidate molecules or substances
 CC that are capable of modulating the transcription of the ABCI gene. They
 CC are used in antisense therapy. Compositions comprising sequences of the
 CC invention are used to treat hypercholesterolaemia and atherosclerosis.
 CC The present sequence is human ABCI gene exon 1A.

SQ Sequence 221 BP; 44 A; 62 C; 73 G; 42 T; 0 other;

Query Match: 16.1%; Score 91; DB 24; Length 221;

Best Local Similarity 100.0%; Pred. No. 1.1e-14;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 GTAATTGCGAGCGAGAGTACGTGGCGCGGGACCGCGAGCGCGACCCCTCTCTC 533

DB 1 GTAATTGCGAGCGAGAGTACGTGGCGCGGGACCGCGAGCGCGACCCCTCTCTC 60

QY 534 CCGGGCTGGCGAGCGCGAGCGCGGAGCTC 564

DB 61 CCGGGCTGGCGAGCGCGAGCGCGGAGCTC 91

RESULT 11

AAH07432

ID AAH07432 standard; cDNA; 736 BP.

XX

AC AAH07432;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:4267.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX

```

XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241599.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises at least 15 nucleotides, where the
XX oligonucleotide comprises at least 15 nucleotides, and the combination
XX of the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;
XX
XX Query Match 15.9%; Score 89.4; DB 22; Length 736;
XX Best Local Similarity 98.9%; Pred. No. 4.3e-14;
XX Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 474 GTAATTGCGAGCAGAGTGTAGTGGCGCGGACCCGAGCGCGAGCCGACCCCTTCTTC 533
XX 1 GTAGTTGCGAGCAGAGTGTAGTGGCGCGGACCCGAGCGCGAGCCGACCCCTTCTTC 60
XX
XX 534 CCGGGCTGGCGAGCAGCGCGGGGGAGCTC 564
XX 61 CCGGGCTGGCGAGCAGCGCGGGGGAGCTC 91
XX
XX RESULT 12
XX AAH18606
XX ID AAH18606 standard: cDNA; 1556 BP.
XX
XX AAH18606;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:18808.
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

```

```

OS Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;
XX
XX Query Match 15.9%; Score 89.4; DB 22; Length 1556;
XX Best Local Similarity 98.9%; Pred. No. 5.4e-14;
XX Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 474 GTAATTGCGAGCAGAGTGTAGTGGCGCGGACCCGAGCGCGAGCCGACCCCTTCTTC 533
XX 1 GTAGTTGCGAGCAGAGTGTAGTGGCGCGGACCCGAGCGCGAGCCGACCCCTTCTTC 60
XX
XX 534 CCGGGCTGGCGAGCAGCGCGGGGGAGCTC 564
XX 61 CCGGGCTGGCGAGCAGCGCGGGGGAGCTC 91
XX
XX RESULT 13
XX AAH04729
XX ID AAH04729 standard: cDNA; 763 BP.
XX
XX AAH04729;
XX
XX 26-JUN-2001 (first entry)

```

```

XX Human cDNA clone (5'-primer) SEQ ID NO:1564.
XX
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX PN EPI074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX
XX PF 28-JUL-2000; 2000EP-0116126.
XX
XX PR 29-JUL-1999; 99JP-0248036.
XX
XX PR 27-AUG-1999; 99JP-0300253.
XX
XX PR 11-JAN-2000; 2000JP-0118776.
XX
XX PR 02-MAY-2000; 2000JP-0183767.
XX
XX PR 09-JUN-2000; 2000JP-0241899.
XX
XX FA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX DR WPI; 2001-318749/34.
XX
XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
XX PT full-length cDNAs defined in the specification, and for the detection
XX PT and/or diagnosis of the abnormality of the proteins encoded by the
XX PT full-length cDNAs -
XX
XX PS Claim 1: SEQ ID 1564; 2537pp + CD ROM; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602
XX CC full-length cDNAs defined in the specification. Where a primer set
XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX CC to the complementary strand of a polynucleotide which comprises one of
XX CC the 5602 nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
XX CC the specification. The primer sets can be used in antisense therapy and
XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX CC represent oligonucleotides, all of which are used in the exemplification
XX CC of the present invention.
XX
XX SQ Sequence 763 bp; 137 A; 205 C; 260 G; 158 T; 3 other;
XX
XX Query Match 15.8%; Score 89; DB 22; Length 763;
XX Best Local Similarity 100.0%; Pred. No. 5.6e-14;
XX Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 476 AATTCCGAGCGAGAGTGAAGTGGGCGGACCCGAGAGCGGACCGACCTTCTCTCC 535
XX Db 1 AATTCCGAGCGAGAGTGAAGTGGGCGGACCCGAGAGCGGACCGACCTTCTCTCC 60
XX
XX QY 536 GGGCTGGCGGCGGAGGCGGGGAGCTC 564
XX Db 61 GGGCTGGCGGCGGAGGCGGGGAGCTC 89
XX
XX RESULT 14
XX AAH17451

```

```

ID AAH17451 standard; cDNA; 1750 BP.
XX
XX AC AAH17451;
XX
XX DT 26-JUN-2001 (first entry)
XX
XX DE Human cDNA sequence SEQ ID NO:16905.
XX
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX OS Homo sapiens.
XX
XX PN EPI074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX PF 28-JUL-2000; 2000EP-0116126.
XX
XX PR 29-JUL-1999; 99JP-0248036.
XX
XX PR 27-AUG-1999; 99JP-0300253.
XX
XX PR 11-JAN-2000; 2000JP-0118776.
XX
XX PR 02-MAY-2000; 2000JP-0183767.
XX
XX PR 09-JUN-2000; 2000JP-0241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX DR WPI; 2001-318749/34.
XX
XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
XX PT full-length cDNAs defined in the specification, and for the detection
XX PT and/or diagnosis of the abnormality of the proteins encoded by the
XX PT full-length cDNAs -
XX
XX PS Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602
XX CC full-length cDNAs defined in the specification. Where a primer set
XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX CC to the complementary strand of a polynucleotide which comprises one of
XX CC the 5602 nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
XX CC the specification. The primer sets can be used in antisense therapy and
XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX CC represent oligonucleotides, all of which are used in the exemplification
XX CC of the present invention.
XX
XX SQ Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;
XX
XX Query Match 15.8%; Score 89; DB 22; Length 1750;
XX Best Local Similarity 100.0%; Pred. No. 7.2e-14;
XX Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 476 AATTCCGAGCGAGAGTGAAGTGGGCGGACCCGAGAGCGGACCGACCTTCTCTCC 535
XX Db 1 AATTCCGAGCGAGAGTGAAGTGGGCGGACCCGAGAGCGGACCGACCTTCTCTCC 60
XX
XX QY 536 GGGCTGGCGGCGGAGGCGGGGAGCTC 564
XX
XX RESULT 14
XX AAH17451

```


GenCore version 5.1.3
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: February 3, 2003, 12:46:48 : Search time 96.1359 Seconds
 (without alignments)
 8245.650 Million cell updates/sec
 Title: US-09-596-141c-3_COPY_1292_1643
 Perfect score: 352
 Sequence: 1 ggaacctaagacacctgtg.....caggggcaggggcgggagctc 352
 Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
 Total number of hits satisfying chosen parameters: 4370478
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :	N_Geneseq_101002.*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352	100.0	1197	24	ABL58400 Human large ATP-bi
2	352	100.0	1643	22	AAF24581 Nucleotide sequenc
3	352	100.0	1643	22	AAF24703 Nucleotide sequenc
4	336	95.5	183999	22	AAF92831 Human ABC1 genomic
5	335.4	95.3	3231	24	AAD37265 Human ABC1 transcr
6	261.4	74.3	2910	6	AAD37267 Human ABC1 transcr
7	119.4	33.9	227	21	AAC09615 Human secreted pro
8	98	27.8	7260	22	AAD21326 Human ATP binding
9	98	27.8	7260	22	AA170315 Human ATP binding

10	91	25.9	221	24	AA037268	Human: ABC1 gene ex	
11	89.4	25.4	736	22	AAH07432	Human: CDNA clone (
12	89.4	25.4	1556	22	AAH18606	Human: CDNA sequenc	
13	89	25.3	763	22	AAH04729	Human: CDNA clone (
14	89	25.3	1750	22	AAH17451	Human: CDNA sequenc	
15	77.2	21.9	7281	22	AAK51683	Human polynucleoti	
16	75.4	21.4	7086	22	ABA09200	Human ABCA1 homolo	
17	75.4	21.4	7086	22	AAK52657	Human polynucleoti	
18	75	21.3	9854	22	AA06121	Human ABC1 DNA seq	
19	67	19.0	10442	22	AA024680	Nucleotide sequenc	
20	67	19.0	10442	22	AA024702	Nucleotide sequenc	
21	60	17.0	10474	22	AA024685	Nucleotide sequenc	
22	60	17.0	10474	22	AA024686	Nucleotide sequenc	
23	60	17.0	10474	22	AA024707	Nucleotide sequenc	
24	60	17.0	10474	22	AA024708	Nucleotide sequenc	
25	38.2	10.9	114955	20	AAK53491	Human adenosine Al	
c	26	37.4	10.6	1599	16	AAQ97701	Rat melanocortin r
27	36.4	10.3	38	22	AA093064	ABCL polymorphism	
28	36.2	10.3	700	22	AAH92242	Human inflammatory	
c	29	35	9.9	486	23	AA081261	DNA encoding novel
c	30	34.8	9.9	10732	21	AA010594	Gene encoding a su
c	31	34.4	9.8	669	23	AA083553	DNA encoding novel
32	34.2	9.7	1610	22	AA080546	Receptor #34 parti	
33	34	9.7	524	22	AA092108	Human CDNA 5'-end	
34	34	9.7	524	22	AAK93547	Human CDNA clone r	
35	34	9.7	1769	22	AAK94827	Human full-length	
36	33.8	9.6	3168	24	ABQ91991	Human NF-kB activa	
37	33.8	9.6	3168	24	ABQ91992	Human NF-kB activa	
c	38	33.2	9.4	695	22	AAH04659	Human CDNA clone (
c	39	33.2	9.4	1235	22	AA080864	Human extracellular
c	40	33.2	9.4	1877	22	AAH77797	Nucleotide sequenc
c	41	33.2	9.4	1877	22	AAI93858	Human stomach can
c	42	33.2	9.4	1877	22	AAH15566	Human CDNA sequenc
c	43	33	510	22	AAK68466	Human digestive sy	
c	44	33	785	24	ABQ41592	Oligonucleotide fo	
c	45	33	785	24	ABQ41593	Oligonucleotide fo	

ALIGNMENTS

RESULT 1
 ABL58400
 ID ABL58400 standard; DNA; 1197 BP.
 XX
 AC ABL58400;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 DE Human large ATP-binding cassette transporter 1(hABC1) promoter sequence.
 XX
 DE Human large ATP-binding cassette transporter 1; ABC1; promoter;
 KW Human; large ATP-binding cassette transporter 1; ABC1; promoter;
 KW antiarteriosclerotic; gene transfer; transactivator; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200183506-A1.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US13654.
 XX
 PR 25-APR-2000; 2000US-0560372.
 XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 PI Tall AR;
 XX
 DR WPI; 2002-049334/06.
 XX
 PT Novel isolated human large ATP-binding cassette transporter 1 promoter
 PT capable of directing transcription of heterologous coding sequence
 PT positioned downstream to it, useful for expressing foreign DNA in host


```
XX DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
XX atherosclerosis; cholesterol transport; ss.
XX
XX Homo sapiens.
XX
XX WC20C078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 200WO-US16591.
XX
XX 18-JUN-1999; 93US-0140264.
XX
XX 14-SEP-1999; 99US-0153872.
XX
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
XX (UNIV) UNIV WASHINGTON.
XX
XX Lawn RM, Wade D, Oram JF, Garvin M;
XX WPI; 2001-13781/14.
XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX polynucleotides and polypeptides, useful for treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX
XX Disclosure; Page 138-139; 211pp; English.
XX
XX The present sequence represents the 5' flanking region of the human
XX adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
XX resides in cell membranes and utilises ATP hydrolysis to transport a wide
XX variety of substrates across the plasma membrane. ABC1 is a pivotal
XX protein in the apolipoprotein-mediated mobilisation of intracellular
XX cholesterol stores. ABC1 is defective in Tangier disease, a genetic
XX disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
XX gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins
XX are useful for developing pharmaceutical agents for the treatment of
XX heart disease and other disorders associated with hypercholesterolemia
XX and atherosclerosis. The genes are useful for developing screening assays
XX to screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX are also useful as diagnostic indicators of cardiovascular disease and
XX other disorders associated with hypercholesterolemia.
XX
XX Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other:
XX
XX Query Match 100.0%; Score 352; DB 22; Length 1643;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-89;
XX Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGACCTTAAGACACCTGCTGTACCTCCACCCACCCACCCACCTCCGCCCAACTCCC 60
XX
XX 1292 GGACCTTAAGACACCTGCTGTACCTCCACCCACCCACCCACCTCCGCCCAACTCCC 1351
XX
XX 61 TAGATGTGCTGGGGGGCTGACGTGCGCCGTTTAAGGGGGGGGGGGGGGGGGGGGGGG 120
XX
XX 1352 TAGATGTGCTGGGGGGGGCTGACGTGCGCCGTTTAAGGGGGGGGGGGGGGGGGGGGG 1411
XX
XX 121 CTCTCTCTGAGTGACGTGAACTACATAAAGAGGGGGGGGGGGGGGGGGGGGGGGGA 180
XX
XX 1412 CTCTCTCTGAGTGACGTGAACTACATAAAGAGGGGGGGGGGGGGGGGGGGGGGGGA 1471
XX
XX 181 GAGCACAGGCTTTGACCGGATAGTAACCTCTGCGTCTGCGTGGAGCCGGAATCTATAAAGGA 240
XX
XX 1472 GAGCACAGGCTTTGACCGGATAGTAACCTCTGCGTCTGCGTGGAGCCGGAATCTATAAAGGA 1531
XX
XX 241 ACTAGTCCCGGCAAAACCCCGTAATTGGAGCGGAGAGTGTAGTGGGGGGGGGGGGGGGG 300
```

```
Db 1532 ACTAGTCCCGGCAAAACCCCGTAATTGGAGCGGAGAGTGTAGTGGGGGGGGGGGGGGGG 1592
Qy 301 AGCCGAGCGGACCCCTCTCTCCCGGGCTGCGGCAGGAGGAGGCGGGGAGCTC 352
Db 1592 AGCCGAGCGGACCCCTCTCTCCCGGGCTGCGGCAGGAGGAGGCGGGGAGCTC 1643
RESULT 4
AAF92831
ID AAF92831 standard; DNA; 183999 BP.
XX
XX AAF92831;
XX
XX 17-MAY-2001 (first entry)
XX
XX Human ABC1 genomic DNA.
XX
XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
XX Homo sapiens.
XX
XX WO200115676-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 200WO-IB01492.
XX
XX 01-SEP-1999; 99US-0151977.
XX
XX 15-MAR-2000; 2000US-0526193.
XX
XX 23-JUN-2000; 2000US-0213958.
XX
XX (UYBR-) UNIV BRITISH COLOMBIA.
XX (XENO-) XENON GENETICS INC.
XX
XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX WPI; 2001-244356/25.
XX
XX Treating a lower than normal high density lipoprotein-cholesterol
XX (HDL-C) level, a higher than normal triglyceride level, or a
XX cardiovascular disease, by administering a compound that modulates LXR-
XX or RXR-mediated transcriptional activity -
XX
XX Claim 8; Fig 1; 317pp; English.
XX
XX The present invention relates to a method for treating a patient
XX diagnosed as having a lower than normal high density
XX lipoprotein-cholesterol (HDL-C) level, a higher than normal
XX triglyceride level, or a cardiovascular disease, involving
XX administering a compound that modulates LXR- or RXR-mediated
XX transcriptional activity or ABC1 expression or activity.
XX The LXR gene product may be used in an assay to identify
XX compounds useful for the treatment of a disease or condition selected a
XX lower than normal HDL cholesterol level, a higher than normal
XX triglyceride level, and a cardiovascular disease.
XX
XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other:
XX
XX Query Match 95.5%; Score 336; DB 22; Length 183999;
XX Best Local Similarity 98.3%; Pred. No. 3.4e-84;
XX Matches 351; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
XX
XX 1 GGACCTTAAGACACCTGCTGTACCTCCACCCACCCACCTCCGCCCAACTCCC 55
XX
XX 28524 GGACCTTAAGACACCTGCTGTACCTCCANNNCNNCCACCCACCTCCGCCCA 28583
XX
XX 56 CTCCTAGATGTGTGTGGGGGGCTGAGCTCGCCCGTTTAAGGGGGGGGGGGGGGGGG 115
XX
XX 28584 CTCCTAGATGTGTGTGGGGGGCTGAGCTCGCCCGTTTAAGGGGGGGGGGGGGGG 28643
XX
XX 116 ACGTGTCTTCTGCTGAGTGTACTGAACTACATAACAGAGCCCGGGGGGGGGAGG 175
XX
XX
```



```
Db 28644 ACCTGCTTTCTGCTGAGTGACTGAACATACATAAACAGAGGCGCGGAACAGCGGGGGGAGG 28703
QY 176 AGGAGACAGACAGGCTTGGACCGATGATACCTCTCGGCTCGGTGAGCGCGAATCTATAA 235
Db 28704 AGGAGACAGACAGGCTTGGACCGATGATACCTCTCGGCTCGGTGAGCGCGAATCTATAA 28763
QY 236 AAGCAACTAGTCCCGGCAAAAACCCCGTAATTGCGAGCGAGTGAGTGGGGCGGGGACC 295
Db 28764 AAGCAACTAGTCCCGGCAAAAACCCCGTAATTGCGAGCGAGTGAGTGGGGCGGGGACC 29823
QY 296 CGGAGACCGGAGCGGACCTCTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCTC 352
Db 28824 CGGAGACCGGAGCGGACCTCTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCTC 28880

RESULT 5
AAD37265
ID AAD37265 standard; DNA; 3231 BP.
XX
AC AAD37265;
XX
Dn 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #1.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
OS Homo sapiens.
XX
XX W020183746-A2.
XX
PN 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-EP05486.
XX
PR 02-MAY-2000; 2000U5-201280P.
XX
PA (AVET ) AVENTIS PHARMA SA.
XX
PI Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
PI Brewer B, Duverger N, Remaley A, Santamarina-Fojo S;
DR WP; 2002-154404/20.
XX
XX Isolated nucleic acid useful for modifying the ATP-binding cassette 1
XX (ABC1) and screening for candidate modulatory compounds or substances
XX
XX Claim 1; Page 130-131; 152pp; English.
XX
XX The invention relates a nucleic acid which is capable of regulating the
XX transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
XX casual gene for pathologies linked to a dysfunctioning of cholesterol
XX metabolism, including diseases such as atherosclerosis. Polynucleotides
XX of the invention are used to screen candidate molecules or substances
XX that are capable of modulating the transcription of the ABC1 gene. They
XX are used in antisense therapy. Compositions comprising sequences of the
XX invention are used to treat hypercholesterolaemia and atherosclerosis.
XX The present sequence is human ABC1 transcription regulating DNA.
XX
SQ Sequence 3231 BP; 809 A; 773 C; 876 G; 773 T; 0 other;
Query Match 95.3%; Score 335.4; DB 24; Length 3231;
Best Local Similarity 98.3%; Pred. No. 1.9e-84;
Matches 351; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 1 GGACCTAAGACACTGCTGTACCTCCAC-----CCCAACCCCAACCCACCTCCGCCCA 55
Db 2628 GGACCTAAGACACTGCTGTACCTCCACCCCAACCCCAACCCACCTCCGCCCA 2687
Qy 56 CTCCTACATGTGCTGGGGGCTGACGTGCGCCGCTTTAAGGGGCGGGCGGCTCC 115
```

```
Db 2688 CTCCTAGATGTCTCGTGGGCGGCTGAACGTGCGCCGCTTTAAGGGGCGGGCCCGGCTCC 2747
QY 116 ACCTGCTTTCTGCTGAGTGACTGAACATACATAAACAGAGGCGCGGAACAGCGGGGGAGG 175
Db 2748 ACCTGCTTTCTGCTGAGTGACTGAACATACATAAACAGAGGCGCGGAACAGCGGGGGAGG 2807
QY 176 AGGAGACAGACAGGCTTGGACCGATGATACCTCTCGGCTCGGTGAGCGCGAATCTATAA 235
Db 2808 AGGAGACAGACAGGCTTGGACCGATGATACCTCTCGGCTCGGTGAGCGCGAATCTATAA 2867
QY 236 AAGCAACTAGTCCCGGCAAAAACCCCGTAATTGCGAGCGAGTGAGTGGGGCGGGGACC 295
Db 2868 AAGCAACTAGTCCCGGCAAAAACCCCGTAATTGCGAGCGAGTGAGTGGGGCGGGGACC 2927
QY 296 CGGAGACCGGAGCGGACCTCTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCTC 352
Db 2928 CGGAGACCGGAGCGGACCTCTCTCTCCCGGCTGCGGCGAGGCGGCGGAGCTC 2984

RESULT 6
AAD37267
ID AAD37267 standard; DNA; 2910 BP.
XX
AC AAD37267;
XX
Dn 21-AUG-2002 (first entry)
XX
DE Human ABC1 transcription regulatory DNA #3.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX protein_bind 12..23
XX /*tag= a
XX /bound_moiety= "LMO2COM/MYOD"
XX protein_bind 97..107
XX /*tag= b
XX /bound_moiety= "DeltaE1"
XX protein_bind 110..125
XX /*tag= c
XX /bound_moiety= "S8/NKX2.5"
XX protein_bind 196..211
XX /*tag= d
XX /bound_moiety= "S8"
XX protein_bind 228..237
XX /*tag= e
XX /bound_moiety= "GATA"
XX protein_bind 399..410
XX /*tag= f
XX /bound_moiety= "IK2"
XX protein_bind 412..420
XX /*tag= g
XX /bound_moiety= "LYF1"
XX protein_bind 528..539
XX /*tag= h
XX /bound_moiety= "LMO2COM/MYOD/DeltaE1"
XX protein_bind 549..556
XX /*tag= i
XX /bound_moiety= "LYF1"
XX protein_bind 558..568
XX /*tag= j
XX /bound_moiety= "DeltaE1"
XX protein_bind 590..596
XX /*tag= k
XX /bound_moiety= "NKX2.5"
XX protein_bind 608..620
XX /*tag= l
XX /bound_moiety= "NRY/CAA"
XX protein_bind 708..715
XX /*tag= m
```

```
FT protein_bind /bound_moiety= "MZFl"
FT 723..730
FT /*tag= n
FT /bound_moiety= "MZFl"
FT 771..785
FT /*tag= o
FT /bound_moiety= "HfH2/SRY/evil"
FT 803..812
FT /*tag= p
FT /bound_moiety= "CREBFl/VBP"
FT 831..837
FT /*tag= r
FT /bound_moiety= "NKX2.5"
FT 1076..1089
FT /*tag= s
FT /bound_moiety= "GATA"
FT 1173..1188
FT /*tag= t
FT /bound_moiety= "LXRalpha/deltaEF1"
FT 1189..1209
FT /*tag= u
FT /bound_moiety= "DeltaEF1/LXFl/IK2"
FT 1483..1491
FT /*tag= v
FT /bound_moiety= "AP4"
FT 1498..1514
FT /*tag= w
FT /bound_moiety= "LMO2-COM/MYOD/deltaEF1/E47"
FT 1524..1545
FT /*tag= x
FT /bound_moiety= "ZID/deltaEF1"
FT 1597..1607
FT /*tag= y
FT /bound_moiety= "DeltaEF1"
FT 1622..1627
FT /*tag= z
FT /bound_moiety= "PPAR"
FT 1632..1637
FT /*tag= aa
FT /bound_moiety= "PPAR"
FT 1685..1698
FT /*tag= ab
FT /bound_moiety= "USF/NMYC/MYCMAX"
FT 1767..1797
FT /*tag= ac
FT /bound_moiety= "DeltaEF1"
FT 1809..1819
FT /*tag= ad
FT /bound_moiety= "DeltaEF1"
FT 1822..1833
FT /*tag= ae
FT /bound_moiety= "SRY"
FT 1840..1850
FT /*tag= af
FT /bound_moiety= "AP1"
FT 1942..1956
FT /*tag= ag
FT /bound_moiety= "HNF3beta"
FT 1978..1985
FT /*tag= ah
FT /bound_moiety= "NKX2.5"
FT 2008..2016
FT /*tag= ai
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2019..2024
FT /*tag= aj
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2051..2059
FT /*tag= ak
FT /bound_moiety= "GATA"
FT 2104..2111
FT /*tag= al
FT /bound_moiety= "SOX5"
FT 2114..2152
FT /*tag= am
FT /bound_moiety= "SRY/HFH/HNF3beta"
FT 2221..2228
FT /*tag= an
FT /bound_moiety= "MZFl"
FT 2234..2249
FT /*tag= ao
FT /bound_moiety= "IK2/NFkappaB/CREL"
FT 2259..2272
FT /*tag= ap
FT /bound_moiety= "LMO2COM/GATA"
FT 2289..2306
FT /*tag= aq
FT /bound_moiety= "MZFl/SRY"
FT 2313..2318
FT /*tag= ar
FT /bound_moiety= "PPAR"
FT 2321..2326
FT /*tag= as
FT /bound_moiety= "PPAR"
FT 2335..2342
FT /*tag= at
FT /bound_moiety= "MZFl"
FT 2361..2384
FT /*tag= au
FT /bound_moiety= "HNF3beta/SRY/evil"
FT 2426..2433
FT /*tag= av
FT /bound_moiety= "MZFl"
FT 2442..2451
FT /*tag= aw
FT /bound_moiety= "AP4"
FT 2455..2466
FT /*tag= ax
FT /bound_moiety= "SRY"
FT 2491..2498
FT /*tag= ay
FT /bound_moiety= "STAT"
FT 2524..2534
FT /*tag= az
FT /bound_moiety= "STAT/PPAR"
FT 2536..2541
FT /*tag= ba
FT /bound_moiety= "PPAR"
FT 2589..2600
FT /*tag= bb
FT /bound_moiety= "AP2"
FT 2610..2617
FT /*tag= bc
FT /bound_moiety= "MZFl"
FT 2634..2648
FT /*tag= bd
FT /bound_moiety= "LMO2COM/MYOD/E47"
FT 2657..2672
FT /*tag= be
FT /bound_moiety= "RREB1"
FT 2680..2698
FT /*tag= bf
FT /bound_moiety= "MZFl/CMVB"
FT 2728..2740
FT /*tag= bg
FT /bound_moiety= "SP1/GC"
FT 2743..2757
FT /*tag= bh
FT /bound_moiety= "USF/NMYC/ARNT"
FT 2758..2773
FT /*tag= bi
FT /bound_moiety= "NFE2AP1"
FT 2774..2787
FT /*tag= bj
FT /bound_moiety= "XFD1/HFH"
FT 2794..2806
```


CC (coronary heart disease, atherosclerosis) and inflammatory diseases
 CC (psoriasis, lupus erythematosus). The identification of ABC1 as a
 CC transporter for interleukin-beta (IL-1beta) identifies this gene as
 CC a candidate for treatment of inflammatory diseases including rheumatoid
 CC arthritis and septic shock. The present sequence is human ABC1 gene.
 XX
 SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
 Query Match: 27.8%; Score 98; DB 22; Length 7260;
 Best Local Similarity 100.0%; Pred. No. 1.1e-17;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 255 AAACCCCGTAATTCGAGCGAGAGTGAAGTGGCGCGGACCCGAGAGCGCGAGCCGCC 314
 DB 1 AAACCCCGTAATTCGAGCGAGAGTGAAGTGGCGCGGACCCGAGAGCGCGAGCCGCC 50
 QY 315 TTCTCTCCCGGCTCGGCGAGCGAGCGAGCGAGCGAGCTC 352
 DB 61 TTCTCTCCCGGCTCGGCGAGCGAGCGAGCGAGCGAGCTC 98
 RESULT 9
 ID AAI70315
 AC AAI70315;
 XX
 DT 07-JAN-2002 (first entry)
 DE Human ATP binding cassette transporter 1 (ABC1) cDNA.
 XX
 KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;
 KW cholesterol; cardiovascular disease; inflammatory disease;
 KW antiinflammatory; antilipemic; antipsoriatic; dermatological;
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;
 KW polymorphism; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 321..7106
 FT /*tag= a
 FT CDS 501..7106
 FT /*tag= b
 FT /*note= "alternative open reading frame of AAI70314"
 FT variation replace(976,A)
 FT /*tag= c
 FT variation replace(1516,C)
 FT /*tag= d
 FT variation replace(2969,G)
 FT /*tag= e
 FT variation replace(3836,C)
 FT /*tag= f
 XX
 PF EF1136554-AL.
 PN
 PD 26-SEP-2001.
 XX
 PF 24-MAR-2000; 2000EP-0106401.
 XX
 PR 24-MAR-2000; 2000EP-0106401.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Schmitz G, Bodzioch M;
 XX
 DR WPI: 2001-640369/74.
 DR P-PSDB; AAM50228.
 XX
 XX New adenosine triphosphate binding cassette transporter gene
 PT polymorphisms, useful for diagnosing and treating lipid disorders,
 PT cardiovascular diseases and inflammatory diseases
 XX

PS Disclosure; Page 26-28; 4ipp; English.
 XX
 CC The present sequence is that of cDNA encoding the human adenosine
 CC triphosphate (ATP) binding cassette transporter 1 (ABC1) protein
 CC (see AAM50227). The sequence includes an extended open reading
 CC frame (ORF) to that provided by the sequence in AAI70314, using
 CC an alternative ATG codon as initiation codon and thereby adding an
 CC extra 40 N-terminal amino acids to the encoded ABC1 protein (see
 CC AAM50228). The invention provides 4 common polymorphisms in the
 CC ABC1 gene. These were identified by sequencing the ABC1 gene in
 CC different Tangier kindreds. In the variant genes (numbering as in
 CC AAI70314), G is changed to A at position 596, T is changed to C at
 CC position 1136, A is changed to G at position 2589 or G is changed
 CC to C at position 3456, or any combination of these. All of these
 CC polymorphisms alter the amino acid sequence of ABC1 and therefore
 CC may affect its function. The 2 most common polymorphisms (G596A)
 CC and A2589G) are both associated with a decreased in vitro ApoA-1
 CC mediated efflux of cholesterol from mononuclear phagocytes, a
 CC feature typical of Tangier disease. 3 Of the variants (G596A,
 CC A2589G and G3456C) are significantly increased in a population of
 CC men having low high density lipoprotein-cholesterol levels and
 CC established coronary heart disease (CHD) relative to CHD-free
 CC control subjects. The use of the provided ABC1 polymorphisms for
 CC the diagnosis and treatment of lipid disorders, cardiovascular
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus
 CC erythematosus) is claimed. Modulation of ABC1 transcripts or
 CC proteins by antisense or ribozyme technology or RNA decoys is also
 CC claimed.
 XX
 SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
 Query Match: 27.8%; Score 98; DB 22; Length 7260;
 Best Local Similarity 100.0%; Pred. No. 1.1e-17;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 255 AAACCCCGTAATTCGAGCGAGAGTGAAGTGGCGCGGACCCGAGAGCGCGAGCCGCC 314
 DB 1 AAACCCCGTAATTCGAGCGAGAGTGAAGTGGCGCGGACCCGAGAGCGCGAGCCGCC 50
 QY 315 TTCTCTCCCGGCTCGGCGAGCGAGCGAGCGAGCGAGCTC 352
 DB 61 TTCTCTCCCGGCTCGGCGAGCGAGCGAGCGAGCGAGCTC 98
 RESULT 10
 ID AAD37268
 AC AAD37268;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Human ABC1 gene exon 1A.
 XX
 KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
 KW cholesterol metabolism; hypercholesterolaemia; antisense therapy;
 KW exon 1A; ds.
 XX
 CS Homo sapiens.
 PN WO200183746-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 02-MAY-2001; 2001WO-EP05488.
 XX
 PR 02-MAY-2000; 2000US-201280P.
 XX
 PA (AVET) AVENTIS PHARMA SA.
 XX
 PI Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
 PI B-rewer B, Duverger N, Remaley A, Santamarina-Fojo S;
 XX

DR WPI; 2002-154404/20.
 XX
 PT Isolated nucleic acid useful for modifying the ATP-binding cassette 1
 PT (ABC1) and screening for candidate modulatory compounds or substances
 PT
 XX
 PS Claim 4; Page 132; 152pp; English.
 XX
 CC The invention relates a nucleic acid which is capable of regulating the
 CC transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
 CC causal gene for pathologies linked to a dysfunctioning of cholesterol
 CC metabolism, including diseases such as atherosclerosis. Polynucleotides
 CC of the invention are used to screen candidate molecules or substances
 CC that are capable of modulating the transcription of the ABC1 gene. They
 CC are used in antisense therapy. Compositions comprising sequences of the
 CC invention are used to treat hypercholesterolemia and atherosclerosis.
 CC The present sequence is human ABC1 gene exon 1A.
 XX
 SQ Sequence 221 BP; 44 A; 62 C; 73 G; 42 T; 0 other;
 Query Match 25.9%; Score 91; DB 24; Length 221;
 Best Local Similarity 100.0%; Pred. No. 4.3e-16;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 262 GTAATGCGAGCGAGTGTGGGCGCGGACCGCGAGCGAGCGACCGCTTCTCTC 321
 DB 1 GTAATGCGAGCGAGTGTGGGCGCGGACCGCGAGCGAGCGACCGCTTCTCTC 60
 QY 322 CCGGGTGTGGGCGAGCGAGCGGGGGAGCTC 352
 DB 61 CCGGGTGTGGGCGAGCGAGCGGGGGAGCTC 91
 RESULT 11
 AAH07432
 ID AAH07432 standard; cDNA; 736 BP.
 AC AAH07432;
 XX
 XX 26-JUN-2001 (first entry)
 DT
 XX Human cDNA clone (5'-primer) SEQ ID NO:4267.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-0116126.
 PF
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI WPI; 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAB55893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;
 Query Match 25.4%; Score 89.4; DB 22; Length 736;
 Best Local Similarity 98.9%; Pred. No. 1.6e-15;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 262 GTAATGCGAGCGAGTGTGGGCGCGGACCGCGAGCGAGCGACCGCTTCTCTC 321
 DB 1 GTAATGCGAGCGAGTGTGGGCGCGGACCGCGAGCGAGCGACCGCTTCTCTC 60
 QY 322 CCGGGTGTGGGCGAGCGAGCGGGGGAGCTC 352
 DB 61 CCGGGTGTGGGCGAGCGAGCGGGGGAGCTC 91
 RESULT 12
 AAH18606
 ID AAH18606 standard; cDNA; 1556 BP.
 AC AAH18606;
 XX
 XX 26-JUN-2001 (first entry)
 DT
 XX Human cDNA sequence SEQ ID NO:18808.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-0116126.
 PF
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI WPI; 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 8: SEQ ID 18808; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

Query Match 25.4%; Score 89.4; DB 22; Length 1556;
 Best Local Similarity 98.9%; Pred. No. 1.9e-15;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 262 GTAATTGGGAGGAGTGTAGTGGGCGGGACCGCAGAGCCGAGCCCTTCCTC 321

Db 1 GTAGTTGGGAGGAGTGTAGTGGGCGGGACCGCAGAGCCGAGCCCTTCCTC 60

QY 322 CCGGGCTGGGAGGCGGCGGGGGAGGCTC 352

Db 61 CCGGGCTGGGAGGCGGCGGGGGAGGCTC 91

RESULT 13

AAH04729

ID AAH04729 standard; cDNA; 763 BP.

XX AAH04729;

XX AAH04729;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:1564.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai Y, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 1: SEQ ID 1564; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match 25.3%; Score 89; DB 22; Length 763;

Best Local Similarity 100.0%; Pred. No. 2.1e-15;

Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 AATTGGGAGGAGTGTAGTGGGCGGGACCGCAGAGCCGAGCCCTTCCTCC 323

Db 1 AATTGGGAGGAGTGTAGTGGGCGGGACCGCAGAGCCGAGCCCTTCCTCC 60

QY 324 GGGCTGGGAGGCGGCGGGGGAGCTC 352

Db 61 GGGCTGGGAGGCGGCGGGGGAGCTC 89

RESULT 14

AAH17451

ID AAH17451 standard; cDNA; 1750 BP.

XX AAH17451;

XX AAH17451;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:16905.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

```
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki I;
XX DR WPI; 2001-3-8749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
XX PT full-length cDNAs defined in the specification, and for the detection
XX PT and/or diagnosis of the abnormality of the proteins encoded by the
XX PT full-length cDNAs.
XX PS Claim 8; SEQ ID 16905: 2537bp + CD ROM; English.
XX CC The present invention describes primer sets for synthesizing 5602
XX CC full-length cDNAs defined in the specification. Where a primer set
XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX CC to the complementary strand of a polynucleotide which comprises one of
XX CC the 5602 nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
XX CC the specification. The primer sets can be used in antisense therapy and
XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX CC represent oligonucleotides, all of which are used in the exemplification
XX CC of the present invention.
XX SQ Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;
Query Match 25.3%; Score 89; DB 22; Length 1750;
Best Local Similarity 100.0%; Pred. No. 2.6e-15;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 264 AATTGGCGAGGAGTGAAGTGGGCGGGACCCGACAGCGGACCGACCCCTCTCTCC 323
DB 1 AATTGGCGAGGAGTGAAGTGGGCGGGACCCGACCGACCGACCGACCCCTCTCTCC 60
QY 324 GGGCTGGCGGAGGCGAGGGGGGGAGCTC 352
DB 61 GGGCTGGCGGAGGCGAGGGGGGGAGCTC 89
RESULT 15
AAK51683
ID AAK51683 standard; cDNA; 7281 BP.
XX AC AAK51683;
XX DT 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 228.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorder; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200157150-A2.
XX XX
XX PD 09-AUG-2001.
```

```
XX PF 05-FEB-2001; 2001WO-US04098.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX PI Xue AZ, Yang Y, Wejhrman I, Goodrich R;
XX DR WPI; 2001-476283/51.
XX DR P-PSDB; AAK78550.
XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX PT useful in diagnosis and gene therapy.
XX PS Claim 1; Page 1086-1096; 6221pp; English.
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activity/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX CC were missing at the time of publication.
XX SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;
Query Match 21.9%; Score 77.2; DB 22; Length 7281;
Best Local Similarity 86.7%; Pred. No. 7.5e-12;
Matches 85; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 255 AAACCCCGGTAAATTCGAGCGAGAGTGAGTGGGCGGGACCCGAGCGAGCGAGCGACCC 314
DB 22 AATTCCCGGGTCCGACGATTCGTGTGAGTGGGCGGGACCCGAGCGAGCGAGCGACCC 81
QY 315 TTCTCTCCGGGCTGGCGAGGCGAGGGCGGGAGCTC 352
DB 82 TTCTCTCCGGGCTGGCGAGGCGAGGGCGGGGAGCTC 119
Search completed: February 3, 2003, 16:33:17
Job time : 181.279 secs
```

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:45:48 ; Search time 37.9628 Seconds
(without alignments)
8245.650 Million cell updates/sec

Title:	US-99-596-141C-3 COPY	1394	1532
--------	-----------------------	------	------

Perfect score: 139
Sequence: 1 agcccgagctcacatgct.....agccgaatctataaaagaa 139

Scoring table: IDENTITY_NUC
Gapoc 10.0 : Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:

```
Minimum D3 seq length: 0
Maximum D3 seq length: 20
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

Database : N Genesec 101002:*

1: /SID32/cgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID32/cgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID32/cgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID32/cgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID32/cgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID32/cgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID32/cgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID32/cgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID32/cgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID32/cgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID32/cgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID32/cgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID32/cgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID32/cgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID32/cgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID32/cgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID32/cgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID32/cgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID32/cgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID32/cgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID32/cgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID32/cgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID32/cgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID32/cgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	139	100.0	1197	24	AB58400	Human large ATP-bi
2	139	100.0	1643	22	AA24681	Nucleotide sequenc
3	139	100.0	1643	22	AA24703	Nucleotide sequenc
4	139	100.0	183999	22	AA292831	Human ABC1 genomic
5	137.4	98.8	2910	24	AA037267	Human ABC1 transcr
6	137.4	98.8	3231	24	AA037265	Human ABC1 transcr
7	36.4	26.2	38	22	AA93004	ABC1 polymorphism
8	30	21.6	1817	19	AAV04064	Human RNA-binding
9	30	21.6	5868	22	AAK89398	Human digestive sv

ALIGNMENT

RESULT 1

A3L58400

ID ABL58400 standard: DNA: 1197 BP.

AA
AC
AB158400.

XX
DT

Figure 1

KW antiarteriosclerotic; gene

Homosapiens

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX

XX

FF - APR - 2001, 2001W, 031305
YY

PR 28-APR-2000; 2000JS-056037

PA (JYCO) UNIV COLUMBIA NEW

PI Tall AR:

DD FORM 3003-040334 106

	P
"	P
"	P
"	P
-	P
"	P
	P
X X	P

PT capable of directing trans

3
3
3
3
4
3
2
2
2
3
3
3
2
4
3
1
1
0
0
2


```
FT /bound_moiety= "IK2"
FT 412..420
FT /*tag= g
FT /bound_moiety= "LYF1"
FT 528..539
FT /*tag= h
FT /bound_moiety= "LMO2COM/MYOD/DeltaEFl"
FT 549..556
FT /*tag= i
FT /bound_moiety= "LYF1"
FT 558..568
FT /*tag= j
FT /bound_moiety= "DeltaEFl"
FT 593..596
FT /*tag= k
FT /bound_moiety= "LXX2.5"
FT 608..620
FT /*tag= l
FT /bound_moiety= "NFY/CAAC"
FT 708..715
FT /*tag= m
FT /bound_moiety= "MZFl"
FT 723..730
FT /*tag= n
FT /bound_moiety= "MZFl"
FT 771..785
FT /*tag= o
FT /bound_moiety= "HFH2/SRY/EVIL"
FT 803..812
FT /*tag= p
FT /bound_moiety= "CREBp1/VBP"
FT 831..837
FT /*tag= r
FT /bound_moiety= "NKX2.5"
FT 1076..1089
FT /*tag= s
FT /bound_moiety= "GATA"
FT 1173..1188
FT /*tag= t
FT /bound_moiety= "LXRalpha/deltaEFl"
FT 1189..1209
FT /*tag= u
FT /bound_moiety= "DeltaEFl/LYF1/IK2"
FT 1483..1491
FT /*tag= v
FT /bound_moiety= "AP4"
FT 1498..1514
FT /*tag= w
FT /bound_moiety= "LMO2-COM/MYOD/deltaEFl/E47"
FT 1524..1545
FT /*tag= x
FT /bound_moiety= "ZID/deltaEFl"
FT 1597..1607
FT /*tag= y
FT /bound_moiety= "DeltaEFl"
FT 1622..1627
FT /*tag= z
FT /bound_moiety= "PPAR"
FT 1632..1637
FT /*tag= aa
FT /bound_moiety= "PPAR"
FT 1685..1698
FT /*tag= ab
FT /bound_moiety= "USF/NMYC/MYCMAX"
FT 1787..1797
FT /*tag= ac
FT /bound_moiety= "DeltaEFl"
FT 1809..1819
FT /*tag= ad
FT /bound_moiety= "DeltaEFl"
FT 1822..1833
FT /*tag= ae
FT /bound_moiety= "SRY"
FT 1840..1850
FT /*tag= af
FT /bound_moiety= "API"
FT 1942..1956
FT /*tag= ag
FT /bound_moiety= "HNF3beta"
FT 1978..1985
FT /*tag= ah
FT /bound_moiety= "NKX2.5"
FT 2008..2016
FT /*tag= ai
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2019..2024
FT /*tag= aj
FT /bound_moiety= "PPAR/NKX2.5/PPAR"
FT 2051..2059
FT /*tag= ak
FT /bound_moiety= "GATA"
FT 2104..2111
FT /*tag= al
FT /bound_moiety= "SOX5"
FT 2114..2152
FT /*tag= am
FT /bound_moiety= "SYR/HFH/HNF3beta"
FT 2221..2228
FT /*tag= an
FT /bound_moiety= "MZFl"
FT 2234..2249
FT /*tag= ao
FT /bound_moiety= "IK2/NFKappaB/CREL"
FT 2259..2272
FT /*tag= ap
FT /bound_moiety= "LMO2COM/GATA"
FT 2289..2306
FT /*tag= aq
FT /bound_moiety= "MZFl/SRY"
FT 2313..2318
FT /*tag= ar
FT /bound_moiety= "PPAR"
FT 2321..2326
FT /*tag= as
FT /bound_moiety= "PPAR"
FT 2335..2342
FT /*tag= at
FT /bound_moiety= "MZFl"
FT 2361..2384
FT /*tag= au
FT /bound_moiety= "HNF3beta/SRY/EVIL"
FT 2426..2433
FT /*tag= av
FT /bound_moiety= "MZFl"
FT 2442..2451
FT /*tag= aw
FT /bound_moiety= "AP4"
FT 2455..2466
FT /*tag= ax
FT /bound_moiety= "SRY"
FT 2491..2498
FT /*tag= ay
FT /bound_moiety= "STAT"
FT 2524..2534
FT /*tag= az
FT /bound_moiety= "STAT/PPAR"
FT 2536..2541
FT /*tag= ba
FT /bound_moiety= "PPAR"
FT 2589..2600
FT /*tag= bb
FT /bound_moiety= "AP2"
FT 2610..2617
FT /*tag= bc
FT /bound_moiety= "MZFl"
FT 2634..2648
```


RESULT 10
AAK89399/c
-D AAK89399 standard; DNA; 10316 BP.
XX
AC AAK89399;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2975.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WC200155314-A2.
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-USG1324.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-C198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209457.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220563.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232083.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233423.
PR 21-SEP-2000; 2000US-0233424.
PR 25-SEP-2000; 2000US-0233497.
PR 25-SEP-2000; 2000US-0233498.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239035.
PR 13-OCT-2000; 2000US-0239957.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241888.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0248207.
PR 17-NOV-2000; 2000US-0248208.
PR 17-NOV-2000; 2000US-0248209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.

```

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251388.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognozing disorders of the
XX digestive system, particularly cancer and cancer metastases -
XX
XX Disclosure: SEQ ID NO 2975; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention.
XX
XX Sequence 10316 BP; 2525 A; 2093 C; 2534 G; 3164 T; 0 other;
XX
XX
XX Query Match 21.6%; Score 30; DB 22; Length 10316;
XX Best Local Similarity 61.5%; Pred. No. 11;
XX Matches 48; Conservative 0; Mismatches 30; Indels 0; Caps 0;
XX
XX
XX 1 GGGCCCGGCTCAGCGCTTCGTCGAGTGACTGACTGACTACATAACAGAGCGCGGAA 60
XX
XX Db 164 GGGCTGCTCCTCCGCTCCGCGCTGATGGCGCGAAGAGGTGATCAGGGCGAGCGG 105
XX
XX
XX 61 CGGGCGCGGAGGAGGGA 78
XX
XX Db 104 CGGGCGCGGCGCGCGA 87
XX
XX
XX RESULT 11
XX AAF30035/c
XX ID AAF30035 standard; DNA; 81001 BP.
XX
XX AC AAF30035;
XX
XX
XX 23-APR-2001 (first entry)
XX
XX Human apolipoprotein A-IV-related protein (AA4RP) gene.
XX
XX
XX Apolipoprotein A-IV-related protein; AA4RP; human;
XX biallelic marker; lipid metabolism; liver related disorder;
XX obesity; diabetes; coronary heart disease; diagnosis; gene therapy;
XX chromosome 11; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX 5'UTR 10946..12946
XX /*tag= a
XX CDS 13478..15210
XX /*tag= b
XX exon 12947..12958
XX /*tag= c
XX /*number= 1
XX

```

```

FT intron 12959..13469
FT /*tag= d
FT /*number= 1
FT exon 13470..13526
FT /*tag= e
FT /*number= 2
FT intron 13527..13640
FT /*tag= f
FT /*number= 2
FT exon 13641..13752
FT /*tag= g
FT /*number= 3
FT intron 13753..14270
FT /*tag= h
FT /*number= 3
FT exon 14271..15968
FT /*tag= i
FT /*number= 4
FT 3'UTR 15969..17969
FT /*tag= j
FT allele replace(1239,C,T)
FT /*tag= k
FT /*note= "biallelic marker 20-828-311, polymorphic
FT base C or T"
FT allele replace(12347,C,T)
FT /*tag= l
FT /*note= "biallelic marker 17-42-319, polymorphic
FT base C or T"
FT allele replace(15241,C,T)
FT /*tag= m
FT /*note= "biallelic marker 17-41-250, polymorphic
FT base C or T"
FT allele replace(42218,A,G)
FT /*tag= n
FT /*note= "biallelic marker 20-841-149, polymorphic
FT base A or G"
FT allele replace(45442,A,G)
FT /*tag= o
FT /*note= "biallelic marker 20-842-115, polymorphic
FT base A or G"
FT allele replace(77058,C,T)
FT /*tag= p
FT /*note= "biallelic marker 20-853-415, polymorphic
FT base C or T"
FT primer_bind 929..949
FT /*tag= q
FT /*note= "primer 20-828.pu"
FT primer_bind complement(1357..1377)
FT /*tag= r
FT /*note= "primer 20-828.rp"
FT primer_bind 12029..12050
FT /*tag= s
FT /*note= "primer 17-42.pu"
FT primer_bind complement(12581..12603)
FT /*tag= t
FT /*note= "primer 17-42.rp"
FT primer_bind 14992..15012
FT /*tag= u
FT /*note= "primer 17-41.pu"
FT primer_bind complement(15460..15482)
FT /*tag= v
FT /*note= "primer 17-41.rp"
FT primer_bind 42070..42090
FT /*tag= w
FT /*note= "primer 20-841.pu"
FT primer_bind complement(42572..42591)
FT /*tag= x
FT /*note= "primer 20-841.rp"
FT primer_bind 45328..45347
FT /*tag= y
FT /*note= "primer 20-842.pu"
FT primer_bind complement(45863..45883)
FT /*tag= z
FT

```



```
FT CDS 1353..2674
FT /tag= a "RRV R1"
FT /product= "complement (2692...3259)"
FT CDS
FT /tag= b
FT /product= "dihydrofolate reductase"
FT /label= RRV_ORF2
FT /note= "has similarity to Kaposi's sarcoma-associated
FT virus (KSHV) open reading frame (ORF) 2"
FT CDS 3676..5613
FT /tag= c
FT /product= "complement binding protein"
FT /label= RRV_ORF3
FT /note= "has similarity to KSHV ORF4"
FT CDS 6045..9443
FT /tag= d
FT /product= "ssDNA binding protein"
FT /label= RRV_ORF6
FT /note= "has similarity to KSHV ORF6"
FT CDS 9468..11528
FT /tag= e
FT /product= "transport protein"
FT /label= RRV_ORF7
FT /note= "has similarity to KSHV ORF7"
FT CDS 11515..14004
FT /tag= f
FT /product= "glycoprotein B"
FT /label= RRV_ORF8
FT /note= "has similarity to KSHV ORF8"
FT CDS 14122..17166
FT /tag= g
FT /product= "DNA polymerase protein"
FT /label= RRV_ORF9
FT /note= "has similarity to KSHV ORF9"
FT CDS 17261..18511
FT /tag= h
FT /label= RRV_ORF10
FT /note= "has similarity to KSHV ORF10"
FT CDS 18520..19749
FT /tag= i
FT /label= RRV_ORF11
FT /note= "has similarity to KSHV ORF11"
FT CDS
FT /tag= j
FT /product= "RRV R2"
FT /label= RRV_ORF
FT /note= "has similarity to KSHV interleukin (IL)-6 gene"
FT CDS
FT /tag= k
FT /product= "thymidylate synthase"
FT /label= RRV_ORF70
FT /note= "has similarity to KSHV ORF70"
FT CDS
FT /tag= l
FT /product= "RRV R3"
FT /note= "has similarity to KSHV K4 viral MIP gene"
FT CDS 26846..27409
FT /tag= m
FT /product= "Bcl2-homologue"
FT /label= RRV_ORF16
FT /note= "has similarity to KSHV ORF16"
FT CDS
FT /tag= n
FT /label= RRV_ORF17
FT /note= "has similarity to KSHV ORF17"
FT CDS 28998..29897
FT /tag= o
FT /label= RRV_ORF18
FT /note= "has similarity to KSHV ORF18"
FT CDS
FT /tag= p
FT /product= "tegument protein"
FT /label= RRV_ORF19
FT /note= "has similarity to KSHV ORF19"
FT CDS
FT /note= "has similarity to KSHV ORF19"
FT complement (31043...32095)
FT /tag= q
FT /label= RRV_ORF20
FT /note= "has similarity to KSHV ORF20"
FT CDS 32094..33767
FT /tag= r
FT /product= "thymidine kinase"
FT /label= RRV_ORF21
FT /note= "has similarity to KSHV ORF21"
FT CDS 33754..35868
FT /tag= s
FT /product= "glycoprotein H"
FT /label= RRV_ORF22
FT /note= "has similarity to KSHV ORF22"
FT CDS
FT complement (35865...37073)
FT /tag= t
FT /label= RRV_ORF23
FT /note= "has similarity to KSHV ORF23"
FT CDS
FT complement (37123...39321)
FT /tag= u
FT /label= RRV_ORF24
FT /note= "has similarity to KSHV ORF24"
FT CDS 39323..43459
FT /tag= v
FT /product= "major capsid protein"
FT /label= RRV_ORF25
FT /note= "has similarity to KSHV ORF25"
FT CDS 43491..44408
FT /tag= w
FT /product= "capsid protein"
FT /label= RRV_ORF26
FT /note= "has similarity to KSHV ORF26"
FT CDS 44433..45242
FT /tag= x
FT /label= RRV_ORF27
FT /note= "has similarity to KSHV ORF27"
FT CDS 45408..45683
FT /tag= y
FT /label= RRV_ORF28
FT /note= "has similarity to KSHV ORF28"
FT CDS
FT /tag= z
FT /label= RRV_ORF29b
FT /note= "has similarity to KSHV ORF29b"
FT CDS 46505..47135
FT /tag= aa
FT /label= RRV_ORF30
FT /note= "has similarity to KSHV ORF30"
FT CDS 47093..47746
FT /tag= ab
FT /label= RRV_ORF31
FT /note= "has similarity to KSHV ORF31"
FT CDS 47883..49077
FT /tag= ac
FT /label= RRV_ORF32
FT /note= "has similarity to KSHV ORF32"
FT CDS 49049..50059
FT /tag= ad
FT /label= RRV_ORF33
FT /note= "has similarity to KSHV ORF33"
FT CDS
FT complement (49977...50960)
FT /tag= ae
FT /label= RRV_ORF35a
FT /note= "has similarity to KSHV ORF29a"
FT CDS 50959..51942
FT /tag= af
FT /label= RRV_ORF34
FT /note= "has similarity to KSHV ORF34"
FT CDS 51923..52372
FT /tag= ag
FT /label= RRV_ORF35
FT /note= "has similarity to KSHV ORF35"
```

```

FT CDS 52278..53585
FT /tag= ah
FT /product= "kinase"
FT /label= RVV_ORF36
FT /note= "has similarity to KSHV ORF36"
FT CDS 53566..55008
FT /tag= ai
FT /product= "alkaline exonuclease"
FT /label= RVV_ORF37
FT /note= "has similarity to KSHV ORF37"
FT CDS 54963..55172
FT /tag= aj
FT /label= RVV_ORF38
FT /note= "has similarity to KSHV ORF38"
FT complement (55255..56391)
FT /tag= ak
FT /product= "glycoprotein M"
FT /label= RVV_ORF39
FT /note= "has similarity to KSHV ORF39"
FT CDS 56526..57932
FT /tag= al
FT /product= "helicase/primase"
FT /label= RVV_ORF40
FT /note= "has similarity to KSHV ORF40"
FT 57917..59528
FT /tag= am
FT /product= "helicase/primase"
FT /label= RVV_ORF41
FT /note= "has similarity to KSHV ORF41"
FT complement (58525..59343)
FT /tag= an
FT /label= RVV_ORF42
FT /note= "has similarity to KSHV ORF42"
FT complement (59297..61027)
FT /tag= ao
FT /product= "capsid protein"
FT /label= RVV_ORF43
FT /note= "has similarity to KSHV ORF43"
FT
FT Query Match 20.7%; Score 28.8; DB 24; Length 128139;
FT Best Local Similarity 54.8%; Pred. No. 53;
FT Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
FT
QY 10 CTCACGCTGTTTCIGCTGAGTGACTGAACATACATAACAGAGCGCGGGAACGGGGGGG 69
Db 69172 CTTAACGAGTGCTCTACTACTGCGGACCGAATACCCATGCGCTGCCCCCGGGAGTG 69113
QY 70 GAGGAGGAGGACAGCGCTTTGACCGATAGTAAACCTCTGCGGT 113
Db 69112 GATATGGAGATGTGAAGCAGGAGAGATGTGCACCCCGGCGCT 69069

RESULT 15
AAC64754/C
ID AAC64754 standard; DNA; 133719 BP.
XX
XX AAC64754;
XX
XX 28-FEB-2001 (first entry)
XX
XX Macaca mulatta rhadinovirus 17577 (RRV) genome sequence SEQ ID NO:1.
XX
XX Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;
XX genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;
XX IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;
XX cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;
XX lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;
XX splenomegaly; hypergammaglobulinemia; autoimmune haemolytic anaemia;
XX ds.
XX
XX Macaca mulatta rhadinovirus 17577.
XX
XX WO200028040-A2.

```

```

XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-US25260.
XX
XX 06-NOV-1998; 98US-0107507.
XX 20-NOV-1998; 98US-0109409.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Wong SW, Axthelm MK, Searles RP;
XX WPI: 2000-376552/32.
XX
XX New rhesus rhadino virus for producing non-human primate model useful
XX for testing potential treatments and efficacy of the candidate vaccine
XX for conditions associated with RRV infection -
XX
XX Claim 2; Page 83-122; 141pp; English.
XX
XX The present invention describes a novel rhesus macaque rhadinovirus
XX called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the
XX RRV genome sequence, and AAB53213 to AAB53204 represent the proteins
XX encoded by the genome sequence. The present invention also specifically
XX claims the individual open reading frame (ORF) nucleotide sequences from
XX the genome which encode the individual proteins, but these sequences are
XX not given. A non-human animal infected with RRV can be used for testing
XX the efficacy of drug in the treatment of condition associated with
XX infection with RRV such as Kaposi's sarcoma, lymphoproliferative
XX disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly,
XX hypergammaglobulinemia or autoimmune haemolytic anaemia, by
XX administering the drug to a immuno-compromised non-human primate
XX preferably Rhesus macaque monkey obtained by as a result of infection
XX by Simian Immunodeficiency Virus (SIV). RRV is useful for producing
XX non-human primate model for testing potential treatments for conditions
XX associated with RRV infection. It is also useful for testing the
XX efficacy of the candidate vaccine against RRV infection or conditions
XX associated with its infection by administering the vaccine to the
XX subject capable of infection with RRV, inoculating the subject with RRV
XX and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205
XX to AAB53213 represent sequence used in the exemplification of the
XX present invention.
XX
XX SQ Sequence 133719 BP; 32746 A; 35648 C; 34521 G; 30804 I; 0 other;
XX
XX Query Match 20.7%; Score 28.8; DB 21; Length 133719;
XX Best Local Similarity 54.8%; Pred. No. 54;
XX Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
XX
QY 10 CTCACGCTGTTTCIGCTGAGTGACTGAACATACATAACAGAGCGCGGGAACGGGGGGG 69
Db 74752 CTTAACGAGTGCTCTACTACTGCGGACCGAATACCCATGCGCTGCCCCCGGGAGTG 74693
QY 70 GAGGAGGAGGACAGCGCTTTGACCGATAGTAAACCTCTGCGGT 113
Db 74692 GATATGGAGATGTGAAGCAGGAGAGATGTGCACCCCGGCGCT 74649

Search completed: February 3, 2003, 16:39:33
Job time : 317.106 secs

```


Db 2536 GGACTAGAGAGTCTGGCGCCAGCCCGAGCCAGCGCTTCCGCGCTTATAGGCCGC 2595
QY 181 GGGCCCGGGGGGAGGGGAGAGAGCCGGAGCCCTAAGACACCTGCTGTACCTCC 240
Db 2596 GGGCCCGGGGGGAGGGGAGAGAGCCGGAGCCCTAAGACACCTGCTGTACCTCC 2655
QY 241 AC-----CCCGACCCACCCACCTCCCGCCAACTCCCTAGATGTGTCTGGGGGCTGAA 295
Db 2655 ACCCGACCCACCCACCCACCTCCCGCCAACTCCCTAGATGTGTCTGGGGGCTGAA 2715
QY 296 CGTGGCCCGTTTAAAGGGGGGGGCGGCGGCTCCACGCTCTTCTGTGAGTGAACCTA 355
Db 2716 CGTGGCCCGTTTAAAGGGGGGGGCGGCGGCTCCACGCTCTTCTGTGAGTGAACCTA 2775
QY 356 CATAACACAGAGCGGGGAGCGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGT 415
Db 2776 CATAACACAGAGCGGGGAGCGGGGGGGGAGGAGGAGGAGGAGGAGGAGT 2835
QY 416 AACCTCTGGCTGGTGCAGCCGAATCTATAAAGGAAGTGTCCCGGCAAAACCCCGT 475
Db 2836 AACCTCTGGCTGGTGCAGCCGAATCTATAAAGGAAGTGTCCCGGCAAAACCCCGT 2895
QY 476 AATGGAGGAGAGTGTAGTGGGGGGGCGGAGCCCGAGAGCCGAGCCGACCTTCTCC 535
Db 2896 AATGGAGGAGAGTGTAGTGGGGGGGCGGAGCCCGAGAGCCGAGCCGACCTTCTCC 2955
QY 536 GGCTCTGGCGAGGCGAGGGGGGGGAGCTC 564
Db 2956 GGCTCTGGCGAGGCGAGGGGGGGGAGCTC 2984

RESULT 2
US-09-846-456-3
; Sequence 3, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrline
; APPLICANT: Naudin, Laurent
; APPLICANT: Deneffe, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying It
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent version 3.0
; SEQ ID NO 3
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-3

Query Match 80.9%; Score 456.4; DB 10; Length 2893;
Best Local Similarity 98.7%; Pred. No. 2.1e-115;
Matches 472; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
QY 1 GAGCTCTCTCCCGCCATCCCTCCCTCCCGCTGGAGCAAACTACACAGCAAAAAAAT 60
Db 2416 GAGCTCTCTCCCGCCATCCCTCCCTCCCGCTGGAGCAAACTACACAGCAAAAAAAT 2475
QY 61 TCGCAAGCAGGATTTAGAGGAGCAAAATCCACTGCTGCCCTTGGCTCCCGGAGAGT 120
Db 2476 TCGCAAGCAGGATTTAGAGGAGCAAAATCCACTGCTGCCCTTGGCTCCCGGAGAGT 2535

QY 121 GGACTAGAGAGTCTGGCGCCAGCCCGAGCCCGAGCGCTTCCGCGCTTATAGGCCGC 180
Db 2536 GGACTAGAGAGTCTGGCGCCAGCCCGAGCCCGAGCGCTTCCGCGCTTATAGGCCGC 2595
QY 181 GGGCCCGGGGGGAGGGGAGAGAGCCGGAGCCCTAAGACACCTGCTGTACCTCC 240
Db 2596 GGGCCCGGGGGGAGGGGAGAGAGCCGGAGCCCTAAGACACCTGCTGTACCTCC 2655
QY 241 AC-----CCCGACCCACCCACCTCCCGCCAACTCCCTAGATGTGTCTGGGGGCTGAA 295
Db 2655 ACCCGACCCACCCACCCACCTCCCGCCAACTCCCTAGATGTGTCTGGGGGCTGAA 2715
QY 296 CGTGGCCCGTTTAAAGGGGGGGGCGGCGGCTCCACGCTCTTCTGTGAGTGAACCTA 355
Db 2716 CGTGGCCCGTTTAAAGGGGGGGGCGGCGGCTCCACGCTCTTCTGTGAGTGAACCTA 2775
QY 356 CATAACACAGAGCGGGGAGCGGGGGGGGAGGAGGAGGAGGAGGAGGAGT 415
Db 2776 CATAACACAGAGCGGGGAGCGGGGGGGGAGGAGGAGGAGGAGGAGGAGT 2835
QY 416 AACCTCTGGCTGGTGCAGCCGAATCTATAAAGGAAGTGTCCCGGCAAAACCCCG 473
Db 2836 AACCTCTGGCTGGTGCAGCCGAATCTATAAAGGAAGTGTCCCGGCAAAACCCCG 2893
RESULT 3
US-09-846-456-4
; Sequence 4, Application US/09846456
; Patent No. US20020146792A1
; GENERAL INFORMATION:
; APPLICANT: Rosier, Marie
; APPLICANT: Prades, Catherine
; APPLICANT: Lemoine, Cendrline
; APPLICANT: Naudin, Laurent
; APPLICANT: Deneffe, Patrice
; APPLICANT: Duverger, Nicolas
; APPLICANT: Brewer, Bryan
; APPLICANT: Remaley, Alan
; APPLICANT: Fojo, Silvia
; TITLE OF INVENTION: Regulatory Nucleic Acid for the ABC1 Gene, Molecules Modifying
; FILE REFERENCE: 3806.0505
; CURRENT APPLICATION NUMBER: US/09/846,456
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/201,280
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-846-456-4

Query Match 16.1%; Score 91; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 474 GTAAATTCGAGCGAGAGTGTGGGGCGGGAGCCCGAGAGCCGAGCCGACCTTCTC 533
Db 1 GTAAATTCGAGCGAGAGTGTGGGGCGGGAGCCCGAGAGCCGAGCCGACCTTCTC 60
QY 534 CCGGGTTCGCGGAGGCGAGGCGGGGAGCTC 564
Db 61 CCGGGTTCGCGGAGGCGAGGCGGGGAGCTC 91

RESULT 4
US-09-770-689A-3
; Sequence 3, Application US/09770689A
; Patent No. US2002011517A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.

```
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001079
; CURRENT APPLICATION NUMBER: US/09/770,689A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 98865
; TYPE: DNA
; ORGANISM: HUMAN
US-09-770-689A-3

Query Match          6.7%; Score 38; DB 10; Length 98865;
Best Local Similarity 55.2%; Pred. No. 1.7;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 74 ATTAGAGGAGCAATCCACTGGTCCCTTGGCTGCGCGGGAACGTGGACTAGAGATC 133
DB 2100 ATGAGAGACAGTTACACCGCATGAGCAGCTGGCTAACACAGACCTGGCGAGCGAGTG 2159

QY 134 TCGCGCGCAGCCCCGAGCCCCAGCGCTTCCCGCGGCTTTAGCGCGCGCGCGCGCGG 193
DB 2160 CGCGCGGACAGCAGCGGGGTGCGACCGGGCTGGGGCGGAGCGCGCGCGGGCGG 2219

QY 194 GAGAGGGGAGCGAG 207
DB 2220 GGACGGCTCCTCG 2233

RESULT 5
US-09-815-048-3
; Sequence 3, Application US/09815048
; Patent No. US20020137131A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001180
; CURRENT APPLICATION NUMBER: US/09/815,048
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 42450
; TYPE: DNA
; ORGANISM: Human
US-09-815-048-3

Query Match          6.5%; Score 37.4; DB 10; Length 42450;
Best Local Similarity 51.5%; Pred. No. 1.8;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 106 GCGTGGCGGAGAGCTGGACTAGAGATCTCGCGCGAGCCCGAGCGCCAGCGCTTCCCG 165
DB 1537 GGTCTCGGAGCGAAGCTCTAGAGGGCGCGGAGAGGGCCCGCCCGCCCTTCGGCC 1596

QY 166 GCGTCTTAGCGCGCGCGCGCGCGGAGGAGGAGGAGGAGCGCGGACCGCTTAAGACA 225
DB 1597 CCACCCACAGCCCGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1656

QY 226 CCGTCTTACCTCCACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
DB 1657 CCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1703

RESULT 6
US-10-047-542-44
; Sequence 44, Application US/10047542
; Patent No. US20020168367A1
; GENERAL INFORMATION:
```

```
; APPLICANT: LARRICK, JAMES W.
; APPLICANT: WYCOFF, KEITH L.
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
; TITLE OF INVENTION: AND BACTERIAL DISEASES
; FILE REFERENCE: 030905.0004.CIP1
; CURRENT APPLICATION NUMBER: US/10/047,542
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/US01/11932
; PRIOR FILING DATE: 2001-04-28
; PRIOR APPLICATION NUMBER: 60/200,298
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-047-542-44

Query Match          6.6%; Score 37.2; DB 9; Length 1920;
Best Local Similarity 49.5%; Pred. No. 0.67;
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 148 GAGCCAGCGCTTCCCGCGGTCTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 207
DB 508 GAACACCCAGCATGCTGGGACCCCGCGGACGCTGGGGCGCAAGAGGAGGCGACACAG 567

QY 208 ACGCGGAGCCCTAAGACACCTGCIGTACCTCCACCCCGCGCGCGCGCGCGCGCGCG 267
DB 568 AGCTCAGAGAGGCCAACACCCCTCATGACACACGATCTTACAGTCTGCTCGACGCGCGCGCGCGCG 627

QY 268 CTCCTAGATGTGCTGGCGGCGCTGAAGCTCCCGCTTTAAGGGCGCGCGCGCGCGCTCC 327
DB 628 TTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 687

QY 328 ACGTGTCTTCTGCT 341
DB 688 CGGACCATCCAGCT 701

RESULT 7
US-09-799-462A-17
; Sequence 17, Application US/09799462A
; Patent No. US20020160970A1
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; Szalay, Aladar
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
; AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAlliff
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,462A
; FILING DATE: 10-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,682
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: 08/695,191
; FILING DATE: 07-AUG-1996
; APPLICATION NUMBER: 08/682,080
; FILING DATE: 15-JUL-1996
```


Db	1525	CAGGAGCCGCGGCTCTCCCTCCGGTCCCGCGGCCGAGTTC	CGAAGCCCTCC	1456
Qy	256	ACCTCCCCCACTCCCTAGATGTCCTGGGGGCGTGAAGCTCGCCGTTTAAAGGGCG	315	
Db	1485	GGCTCTCCCTGCGGAGCGGGCGCGGGCGGGAATGCGCGGATGAGCTCC	1406	
Qy	316	GGCCCCGGCTCCAC	329	
Db	1405	GGAGCCGCTCCC	1392	

```

RESULT 13
US-09-834-975-1012/c
; Sequence 1012, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Haffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2003-04-13
; PRIOR APPLICATION NUMBER: 63/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1012
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1797)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-1012

```

RESULT 14
US-09-834-975-1024/C
; Sequence 1024, Application US/09834975
; Patent No. US20020110813A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

```

; TITLE OF INVENTION:  OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834, 975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1024
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (...)(1797)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-1024

```

```

RESULT 15
US-09-764-869-2109/c
; Sequence 2109, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PAKM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2109
; LENGTH: 550C
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-869-2109

```

Query Match:	6.1%;	Score 34.6;	DB 10;	Length 1550C;
Best Local Similarity	55.4%;	Pred. No. 7.3;		
Matches	67; Conservative	0; Mismatches	54; Indels	0; Gaps
QY	131	GTCTGGGCGCAGCCCGAGCCAGCCGCTTCCCGCGGTCTATAGCGGGCGGGGCCCGGC	190	
Dd	504	GTCGCTCGCGTCGTCCTTGCTACCTGTGTACTGCGCGGGGSCCGCGGGGCCCATGG	445	
QY	191	GGGGGAAGGGAGCAGACCGCGGACCCTTAAGACACCTGTGTATACCTTCCACCCCACCC	250	
Dd	444	GGCGAAGCGGGCGCCCCCGGGGCCCCCAGCCCTCTCGCGGACCCGGTCCCGGAGCG	385	
QY	251	C	251	
Dd	384	C	384	

Tue Feb 4 09:39:41 2003

us-09-596-141c-3_copy_1080_1643.rnpb

Page 7

Search completed: February 3, 2003, 16:30:02
Job time : 130.834 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 ; Search time 4386.17 Seconds
(without alignments)
10165.006 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1_1532
Perfect score: 1532
Sequence: 1 gaattccttggtggtggtc.....ggccgaatctctaaaagaza 1532

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GccEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vt.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vt.*

29: em_vi.*

30: em_hgt_hum.*

31: em_hgt_inv.*

32: em_hgt_other.*

33: em_hgt_mus.*

34: em_hgt_pin.*

35: em_hgt_rtd.*

36: em_hgt_nam.*

37: em_hgt_vrt.*

38: em_sy.*

39: em_hgt_hum.*

40: em_hgt_mus.*

41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1532	100.0	1843	6	AX060715	Sequence
2	1532	100.0	1843	6	AX060894	Sequence
c 3	1433.8	93.6	96717	9	AL359182	Human DNA
4	1387.6	90.6	183999	6	AX092589	Sequence
5	1374.4	89.7	2893	6	AX351031	Sequence
6	1374.4	89.7	3231	6	AX351039	Sequence
7	1307	85.3	149034	9	AF275948	Homo sapi
8	1138.2	74.3	201144	9	AF287262	Homo sapi
9	896	58.5	175064	2	AC012230	Homo sapi
10	896	58.4	1167	9	HS4252201	Homo sapi
11	893.4	58.3	1167	9	AF258623S1	Homo sapi
12	408.8	26.7	69570	2	AC021246	Homo sapi
c 13	171.8	11.2	69570	2	AC021246	Homo sapi
c 14	135.6	8.9	90688	2	AC021345	Homo sapi
c 15	128.8	8.4	186889	2	AL807243	Mus muscu
c 16	128.8	8.4	278572	10	AF287263	Mus muscu
c 17	126.6	8.3	145833	2	AC125837	Rattus no
c 18	102.2	6.7	2893	6	AX351031	Sequence
c 19	102.2	6.7	3231	6	AX351039	Sequence
c 20	99.4	6.5	96717	9	AL359182	Human DNA
c 21	97.8	6.4	149034	9	AF275948	Homo sapi
c 22	97	6.3	1843	6	AX060715	Sequence
c 23	97	6.3	1843	6	AX060894	Sequence
c 24	92.6	6.0	183999	6	AX092589	Sequence
c 25	69.2	4.5	201144	9	AF287262	Homo sapi
c 26	62	4.0	597	9	AF258627	Homo sapi
27	55.6	3.6	191720	2	AC109982	Rattus no
28	52.6	3.4	125020	9	AF429315	Homo sapi
c 29	52.2	3.4	125020	9	AF429315	Homo sapi
c 30	49.6	3.2	96136	2	AC097674	Rattus no
c 31	49.4	3.2	173127	2	AC131142	Rattus no
c 32	49	3.2	65780	2	AC111741	Rattus no
33	48.6	3.2	13645	2	AC108454	Homo sapi
34	48.4	3.2	137136	2	AC108905	Homo sapi
35	48.4	3.2	141114	2	AC108903	Felis cat
c 36	48.4	3.2	206789	2	AC109180	Mus muscu
c 37	47.6	3.1	151975	2	AC109403	Rattus no
c 38	47.6	3.1	190343	2	AC097794	Rattus no
c 39	47.4	3.1	101492	2	AC116065	Rattus no
c 40	47.4	3.1	180461	2	AC106565	Rattus no
c 41	47	3.1	106983	9	AC108677	Rattus no
c 42	47	3.1	110000	2	AC096324_0	Homo sapi
c 43	47	3.1	167886	2	AC126316	Rattus no
c 44	46.8	3.1	2870	9	HS4800882	Homo sapi
c 45	46.8	3.1	169799	2	AC098129	Rattus no

ALIGNMENTS

RESULT 1
LOCUS AX060715
DEFINITION Sequence 3 from Patent WO0078972.
ACCESSION AX060715
VERSION AX060715.1 GI:12406104
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1643)
AUTHORS Lawn,R.M., Wade,D. and Garvin,M.
TITLE Regulation with binding cassette transporter protein abcl
JOURNAL Patent: WO 0078972-A 3 28-DEC-2000;

Db	1	GAATCTCTGCTGGTGGCTCCACATGCACCTCCAGGGCTCTCTGGCTCTCTATGGGTC	60
QY	61	TGTCCTCAGTGTGATAGAACCACTGA-GTGAGTACCTGGCTTCAGCTGGCCTGGAGA	120
Db	61	TGTCCTCAGTGTGATAGAACCACTGATGTAGTACCTTGGCTTCAGGCTGGCCTGGAGA	120
QY	121	TCCTGTGTACTGTAGCATGAGAGGGGCTGTGCAGCTGAATGCTGTATGCAGGTGGTGGG	180
Db	121	TCCTG-TGACTGTAGCATGAGGGGCTGTGCAGTGAA-TGCTGTATGCAGGTGGTGGG	180
QY	181	AGTTC-TGGAATATGATGGAGCTGGAGTGGAGAGAGATAGGCTTGGGCGAGCTCTCTC	240
Db	181	AGTTC-TGGAA-TATGATGGAGCTGGAGTGGGAAGAGAATAGGCTTGGGCGAGCTCTCTC	240
QY	241	AUGCACCTCAATCTCGCCCAAAAC-CAGGICAAACTGTGAAGAGTCTAAATGTGAATCTG	300
Db	241	ATGCCACCTCATCTCGCCCAAAACTCAGGICAAACTGTGAAGAGTCTAAATGTGAATCTG	300
QY	301	CCCTCAAGTGGCTACAAAGGTATCTTTGTCAAGTAGGAGACCTTGTGGCCCTCCACGT	360
Db	301	CCCTCAAGTGGCTACAAAGGTATCTTTGTCAAGTAGGAGACCTTGTGGCCCTCCACGT	360
QY	361	GCACTTCCAGGGCTCTTGGGCTCTTCTACGGTCTGTCTGGTGTCTTCTA-TGAATCTC	420
Db	361	GCACTTCCAGGGCTCTTGGGCTCTTCTAGGGTCTCTCTGGTGTCTTCTA-TGAATCTC	420
QY	421	CCTTCAGGCGAGATTCATATTTAGACTCTTCACAGTTTGACCTCAGTTTGGCCAGAA-TA	480
Db	421	CCTTCAGGCGAGATTCATATTTAGACTCTTCACAGTTTGACCTCAGTTTGGCCAGAA-TA	480
QY	481	AGTGACATTTAG-TTGTGGCTTGATGAATGACPTAAATATTTAGACATATGCTGTGTA	540
Db	481	AGTGACATTTAGTTGTGGCTTGATGAATGACPTAAATATTTAGACATATGCTGTGTA	540
QY	541	GGCTTCGATCTACTCTTGGCTTTTCTTTTGGCCCTCCAGTGTTTTGGGTAGTTTGGCT	600
Db	541	GGCTTCGATCTACTCTTGGCTTTTCTTTTGGCCCTCCAGTGTTTTGGGTAGTTTGGCT	600
QY	601	CCCCCTACAGCCAAAGGCAACAGATAAGTTGSGAGTCTGGAGTGGCTACATATTTTAC	660
Db	601	CCCCCTACAGCCAAAGGCAACAGATAAGTTGSGAGTCTGGAGTGGCTACATATTTTAC	660
QY	661	ACGACTGCATTTCTGTGGCTGCACITCACTCAATGATATCAAACTAAATCAAGTCCCTGTG	720
Db	661	ACGACTGCATTTCTGTGGCTGCACITCACTCAATGATATCAAACTAAATCAAGTCCCTGTG	720
QY	721	TTTTTATCACGGGAGCTGATCAATATAA-TGAATTAAGAGGGGCTGGTCCCATATTG	780
Db	721	TTTTTATCACGGGAGCTGATCAATATAA-TGAATTAAGAGGGGCTGGTCCCATATTG	780
QY	781	TTCTGTGTTTTTGTGTGTTTGTTCCTTTTGTTTTGTGGCCCTCTCTCTCAATTT	840
Db	781	TTCTGTGTTTGTGTGTTTGTTCCTTTTGTTTTGTGGCCCTCTCTCTCAATTT	840
QY	841	ATGAAGAGAGCAGTAAGATGTTCTCTCGGCTCTCTGAGGGACCTGGGAGCTCAGGC	900
Db	841	ATGAAGAGAGCAGTAAGATGTTCTCTCGGCTCTCTGAGGGACCTGGGAGCTCAGGC	900
QY	901	TGGGAATCTCCAAAGCAGTAGTGCGCTATCAAAATCAAAAGTCCAGTTTGTGGGGGA	960
Db	901	TGGGAATCTCCAAAGCAGTAGTGCGCTATCAAAATCAAAAGTCCAGTTTGTGGGGGA	960
QY	961	AAACAAAAGCAGCCCATACCACAGGACTCTCGCCCTTCCCTCCACCCAGGCTAGGCC	1020
Db	961	AAACAAAAGCAGCCCATACCACAGGACTCTCGCCCTTCCCTCCACCCAGGCTAGGCC	1020
QY	1021	TTTCAAGGAAACAAAGACAAGCAAAATGATTGGCTCTCTGAGGAGATTACGCCCTAG	1080
Db	1021	TTTCAAGGAAACAAAGACAAGCAAAATGATTGGCTCTCTGAGGAGATTACGCCCTAG	1080
QY	1081	AGCTCTCTCTCCCCAAATCCCTCCCTCCGGCTGAGGAAACTAACAAAGGAAAAAAATTT	1140
Db	1081	AGCTCTCTCTCCCCAAATCCCTCCCTCCGGCTGAGGAAACTAACAAAGGAAAAAAATTT	1140

QY	1141	CGGAAACAGAGATTAGAGGAAGCAAAATCCACTGTGCCCTTGGCTGCGGGAACGTG	1200
DB	1141	CGGAAACAGAGATTAGAGGAAGCAAAATCCACTGTGCCCTTGGCTGCGGGAACGTG	1200
QY	1201	GACTAGACAGTCTCGGGGCGAGCCCGGAGCCAGCGCTTCGGCGGCTCTTAGCGCGGCG	1250
DB	1201	GACTAGACAGTCTCGGGGCGAGCCCGGAGCCAGCGCTTCGGCGGCTCTTAGCGCGGCG	1250
QY	1261	GCGCGGCGGGGGAAGGGGACGAGACCGGGACCTTAAGACACCTGCTCTGACCCCTCCA	1320
DB	1261	GCGCGGCGGGGGAAGGGGACGAGACCGGGACCTTAAGACACCTGCTCTGACCCCTCCA	1320
QY	1321	CCCCACCCACACCTCCCTCCCGCAACGCCCTAGATGTGTCGTGGCGGTGTGAACGTGC	1380
DB	1321	CCCCACCCACACCTCCCTCCCGCAACGCCCTAGATGTGTCGTGGCGGTGTGAACGTGC	1380
QY	1381	CCGTTTAAGGGGCGCGCGCGCTCCACGCTCTTCGCTGAGTGACTGACTACATAAA	1440
DB	1381	CCGTTTAAGGGGCGCGCGCGCTCCACGCTCTTCGCTGAGTGACTGACTACATAAA	1440
QY	1441	CAGAGCCCGGAACGCGCGGGGAGGAGAGACACAGGCTTTGACCGATGATGAACCTC	1500
DB	1441	CAGAGCCCGGAACGCGCGGGGAGGAGAGACACAGGCTTTGACCGATGATGAACCTC	1500
QY	1501	TGCGCTCGGTGACCGCGAATCTATAAAGGAA	1532
DB	1501	TGCGCTCGGTGACCGCGAATCTATAAAGGAA	1532

RESULT 3
 AL359182/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AL359182 Human DNA sequence from clone RP11-217B7 on chromosome 9, complete sequence.
 AL359182
 AL359182.20 GI:18151453
 HIG.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 - (bases 1 to 36717)
 SkUc.C
 Direct Submission
 Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB0 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Jan 15, 2002 this sequence version replaced gi:18121468.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
 RP11-217B7 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECT03: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-217B7. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RP11-217B7 is at 96717 in this sequence. The true left end of clone RP11-122F10 is at 72980 in this sequence. The true right end of clone RP11-31J20 is at 2000 in this sequence.

FEATURES

source

Location/Qualifiers

1..96717

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/clone="RP11-217B7"

/clone_lib="RP11-11.1"

3238..3278

/note="Sequence from AF275943 sequenced by National Institutes of Health, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA."

misc_feature

84249..84273

/note="Sequence from overlapping clone RP11-122F10 (AC026643). Assembly confirmed by restriction digest."

misc_feature

92050..92153

/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

misc_feature

92411..92557

/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

BASE COUNT

27673 a 21138 c 20380 g 27526 t

Query Match

Best Local Similarity 93.6%; Score 1433.8; DB 9; Length 96717;

Matches 1513; Conservative 0; Mismatches 2; Indels 8; Gaps 7;

QY 16 GGCTCCACATGCACTTCCAGGGCGCTGG--CTCTCTATGGGTCTGCCCAGTGT 73
 DB 18798 GCCTCCACATGCACTTCCAGGGCGCTGTTGGGCTCTCTATGGGTCTGCCCAGTGT 18739
 QY 74 GATAGAACCACTGATGTGATGACTGGGTGTGAG-CGTGGCCCTGGAGATCTCTTGACTG 132
 DB 18738 GATAGAACCACTGATGTGATGACTGGGTGTGAGCGGTGGCCCTGGAGATCTCTTGACTG 18679
 QY 133 TAGCATGGAGGGGGCTTGT-CAGCTGAATCTCTGTATGAGAGTGGTGGGAGTCTCGAAT 131
 DB 18678 TAGCATGGAGGGGGCTTGTGACGTGAATCTCTGTATGAGAGTGGTGGGAGTCTCGAAT 18619
 QY 192 ATGATGGAGCTGGAGTGGGAGAGAGTGGCTTGGGCGAGCTCTCTCATGCCAAGTCA 251
 DB 18618 ATGATGGAGCTGGAGTGGGAGAGAGTGGCTTGGGCGAGCTCTCTCATGCCAAGTCA 18559
 QY 252 TTCTGCCCAAACTCAGTCAAACTGTGAAGAGTCAAAATGTGAATCTGCCCTCAAGGT 311
 DB 18558 TTCTGCCCAAACTCAGTCAAACTGTGAAGAGTCAAAATGTGAATCTGCCCTCAAGGT 18499
 QY 312 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCCCTCCAGCTCCAGTCCAGG 371
 DB 18498 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCCCTCCAGCTCCAGTCCAGG 18439
 QY 372 GCTGCTT-GGCTCTCTACGGGTCTGCTGAGTCTTCTATGAATCT-CCCTTCAGGG 429
 DB 18438 GCTGCTTGGCTCTCTACGGGTCTGCTGAGTCTTCTATGAATCT-CCCTTCAGGG 18379
 QY 430 CAGATTCAATATTAGACTCTTACACTTTCACACTTGTGAGTGTGGCCAGATAAGGTGACAT 489
 DB 18378 CAGATTCAATATTAGACTCTTACACTTTCACACTTGTGAGTGTGGCCAGATAAGGTGACAT 18319
 QY 490 TTAGTTTGTGGCTTGAATGACCTTAATATTAGACATATGGTGTAGCCCTGCAT 549
 DB 18318 TTAGTTTGTGGCTTGAATGACCTTAATATTAGACATATGGTGTAGCCCTGCAT 18259

RESULT 4

AX092569

LOCUS

AX092599

183999 bp

DNA

linear

FAT 21-MAR-2001


```

110646..110666,112124..112196,113183..113385,
115053..115101,115297..115470,116806..116954,
118604..118728,119864..120062,123094..123283,
124804..124898,126208..126240,127363..127468,
128943..129017,129539..129708,130939..131116,
133114..133229,133341..133485,134526..134649,
135736..135865,136130..136250,137038..137100,
138008..138114,140469..140610,140982..141116,
142060..142163,142646..142738,143397..143640,
144381..1448034)
/gene="ABCA1"
/product="ABCA1"
join(1454..1674,25831..25923)
/gene="ABCA1"
2406..2678
/rpt_family="Alu"
repeat_region
/rpt_family="Alu"
4240..4509
/rpt_family="Alu"
repeat_region
4998..5270
/rpt_family="Alu"
repeat_region
5563..5839
/rpt_family="Alu"
repeat_region
6615..6877
/rpt_family="Alu"
repeat_region
8800..9046
/rpt_family="Alu"
repeat_region
9998..10279
/rpt_family="Alu"
repeat_region
11859..12133
/rpt_family="Alu"
repeat_region
12810..12902
/rpt_family="Alu"
repeat_region
15220..15399
/rpt_family="Alu"
repeat_region
15508..15787
/rpt_family="Alu"
repeat_region
18600..18860
/rpt_family="Alu"
repeat_region
20950..21206
/rpt_family="Alu"
repeat_region
25038..25314
/rpt_family="Alu"
repeat_region
CDS
join(25924..25989,40385..40478,45012..45153,46423..46541,
67729..67850,70831..71007,83960..84052,89010..89250,
91962..92101,92433..92549,96758..96955,97702..97907,
98428..98604,100391..100613,102360..102581,103642..103846,
104951..105064,106862..107035,108023..108154,
109460..109602,109806..109943,110646..110866,
112124..112196,113183..113385,115053..115101,
115297..115410,116806..116954,118604..118728,
119964..120062,123094..123283,124804..124898,
126208..126240,127363..127468,128943..129017,
129539..129708,130939..131116,133114..133229,
133341..133485,134526..134649,135736..135865,
136130..136250,137038..137100,138008..138114,
140469..140610,140982..141116,142060..142163,
142646..142738,143397..143640,144381..144721),
/gene="ABCA1"
/codon_start=1
/product="ABCA1"
/protein_id="AAF56276.1"
/db_xref="GI:9247086"
/translation="MACPQLRLLLWKNLFRRTQTCOLLLEAVWPLFIPILISVRL
SYRPEQHECHFPKAMPAGTLPWVGITICANNPCFRYPPTCEAPGVGNKSTV
APLPSDARLLLYNOKTSMKDMKVLRTLOQIKSSNKLQDFLYDNETFSGFLYH
NLSLPKTVDKMLRADVILEKVFLQGLHLTSCNGSKSEMIQLQDQVEVSELCGLP
REKLAARVLRNSMDILKPTLRLNSTSPFPELAEATKLLHSLGTZAQELFSMR
SWSDRQREVFLLTNVSSSSSTIQIQAIVSRVCGHPGCGKLKSLNWDNNYKALF
GGNGTEDEATFFYDNTTPYCNLDMLKNLESSPLSLIKWALKPLLVGKLYTPDIPAT
ROYMAEVNKTFOELAVFDLGMWEELSPKIWTFWENSQENDLYMLLSDSDHFWNE
QQLDGLDWTAGDIYAFILAKPEFDVQSSNGSYITWREAFNETNQAIRTISRMECVNLN
KLEPLATEVWLNKSMELDERKEWAGIVFTGTFGSIELPHVYKIRMDIDNVERT
NKIKDGWDPGDPDPEDMEYVWGGFAYLQDVVEQAIIRVLGTGKTKGYVMOQMPY

```

```

PCYVDDIELVMSSEPLEMTLAWIVAVIILKIGIYVEKEARLKETMRIMGLDLSILW
FSNFISSLILPLLSAGLLVLLKLGILLPLSDPSVFFVFSFAVAVTILQCFLLSTLF
SRANLAACGGIIFFTLLPYLVCWMDYVGTFLKIFAXLLSPVAFGFGCEYFALFE
EOGIGVQMDNLFPSEVEDGFNLTYSMMLEDTFLYGVKWTWYIYAVFGQYGIPIREW
YFCTKYWFGEESDEKSHPSNOKRISEICMESEPHLKLGVSIQNLVKYVRQGMKV
AVDGLALNVEGOITISFLGNGAGKTTMSILTLGLTPPTISGTAVILQKDIRSEKSTIR
ONLQVCPQHNVLFDMLAVESHIWFAVLKGLSEKHVAEMOMALDYLSPSSKLKSKT
SOLSGMQRKLSVALAFVSGSKVILDEPTAGVDPPYRGIWELLILAYRGRITLLST
HMDEADVLGRIATIIISHGLCCVSSGLFKNLQGTGYITLVKKNVSSLSLSCRNSS
STVSYLKEDSVSQSSDAGLGSJESDITLIDVSAISNLIIRKHSARLVEDIGHEL
TYVLYPAKKEGAFVELFHEDDRLSDIGISSYISSETLIEEIPFKVAESGVDAET'S
DGLTPARRNRRAFGDKSCLEPPTEDDAADPNDSDIDPESRETLLSGMDGKGSYQVK
GWKLTQOQFVALLWKRLLIARSRKGFQAQIVLPAVGVIALVPSLIVPFGKPSLE
LOPMXNEQYTFYSNDAPEDGTLELLNALTKDPGCTGRCHGNPDPDFCQAGEEW
TTAPVQIIMDLFONGNWTQNPACCCSDKIKMLPYCPPGAGLPPPPQKQNTA
DILQDLGRNISDYLVKTYVOITIAKSLKNIWNEFRYGFSGVSNQALPPSPQYN
DAXKMKHLKLAKDSSADRLNSLGRFMTGLDRNNVKYNNKNGWHAISSEFLNIN
NAILRANLQGENPSHYGIYAFNHLNLTQQLSEVAXMTTSDVLVSIQVIFAMSV
PASTVFLIOERSVSKAKHLOFISGVKPIYVLSNFWMDMNVVVPATVILIECFQOQ
KSYVSTNLPLVALILLLYGWSITPLMYPASTVFKIPSTAYVLTSLNVLIGINGSVA
TFVLELFDKLNINDILKSVFLIFPHFCGLGRGLDMVKNQAMADALERFGNRFVS
PLSWDLVGRNLFAMAVEGVFFLTIVLIQYRFFIRPRVNAKLSPLNDEDEDVREHO
RILDGGQNDILEIKELTKYRRKRKPAVIRICVGPFGCGLLGVNGAGKSTIFKH
LTGDTVTRGDALFNXNSIISNIHEVQNMGYCPQFDAILTELLTGREHVEFFALLRGV
PEKVGKVGWAIKRLGLVYKERYAGNYSGNKRKLSTAMALITGGVPPVFLDPTTG
MDKARBLNCLSVYKSGSVLTSSEKECEALATRKALWNGRFLGSGVHLK
NRFCDGYTVVRIAGSPDLKPQVDFGLAPPGVVXKEKHNMLQYQLPSSLSSYLK
PSILSQSKRHLHEDYSVQSTLQDQVFNFAKQSDDDHLDLSLHKNOIVDVAVUT
SFLQDERVKESYV"
29916..30180
/rpt_family="Alu"
31050..31314
/rpt_family="Alu"
31692..31973
/rpt_family="Alu"
36167..36440
/rpt_family="Alu"
37617..37892
/rpt_family="Alu"
37999..38166
/rpt_family="Alu"
38172..38269
/rpt_family="Alu"
38995..39271
/rpt_family="Alu"
41547..41565
/gene="ABCA1"
/note="A009X28"
41593..41835
/rpt_family="Alu"
41657..41679
/gene="ABCA1"
/note="A009X28"
42320..42584
/rpt_family="Alu"
44349..44626
/rpt_family="Alu"
48473..48669
/rpt_family="Alu"
50547..50614
/rpt_family="Alu"
54285..54418
/rpt_family="Alu"
54849..54988
/rpt_family="Alu"
55029..55226
/rpt_family="Alu"
55297..55344
/rpt_family="Alu"
58501..58935
/note="LIR"
51246..61489
/rpt_family="HERV"

```

```

/rpt_family="Alu"
repeat_region 62812..63077
/rpt_family="Alu"
repeat_region 54082..64306
/rpt_family="Alu"
repeat_region 67123..67392
/rpt_family="Alu"
repeat_region 68499..68731
/rpt_family="Alu"
repeat_region 69481..69760
/rpt_family="Alu"
repeat_region 71998..72233
/rpt_family="Alu"
repeat_region 72244..72501

Query Match      85.3%; Score 1307; DB 9; Length 149034;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 5; Indels 21; Gaps 8;

QY 113 CTTGGAGATCCTTTGACGTAGCATGAGGAGGGGCTGTG-CAGCTGAATGTCTGTATGCA 171
DB 1 CTTGGAGATCCTTTGACGTAGCATGAGGAGGGGCTTTGTGCACTGAA-TGTCTGCAATGCA 60

QY 172 GGTGCTGGGAGTCTCGAATATGATGGAGCTGGAGCTGGAGAGAGAGTACGCTTGGGGC 231
DB 61 GGTGCTGGGAGTCTCGAATATGATGGAGCTGGAGCTGGAGAGAGAGTACGCTTGGGGC 120

QY 232 AGCTCTCTCAIGCACCTCAITCTGCCAAACTCAGGTCAAACTGTGAAGATCTAAAT 291
DB 121 AGCTCTCTCAIGCACCTCAITCTGCCAAACTCAGGTCAAACTGTGAAGATCTAAAT 180

QY 292 GTGAATCTGCCCTTCAAGTGGCTACAAGAGTATCTTTGTCAAGTAGAGAGACCTTGTGG 351
DB 181 GTGAATCTGCCCTTCAAGTGGCTACAAGAGTATCTTTGTCAAGTAGAGAGACCTTGTGG 240

QY 352 CTTCCAGCTGCATCTCCAGGCTCGCTT-GGCTCTCTACGGGTCTGCTCAGTCTTC 410
DB 241 CTTCCAGCTGCATCTCCAGGCTCGCTTGGGCTCTCTACGGGTCTGCTCAGTCTTC 300

QY 411 TATGAATCT-CCCTTCAGGCGAGATTCATATTTAGACTCTTTCACAGTTTGACCTGAGTTT 469
DB 301 TATGAATCTGTCCTTCAGGCGAGATTCATATTTAGACTCTTTCACAGTTTGACCTGAGTTT 360

QY 470 TGGCCAGATTAAGTGCACATTTAGTTTGTGTGCTTTCATGATCACTTAATATTAGACA 529
DB 361 TGGCCAGATTAAGTGCACATTTAGTTTGTGTGCTTTCATGATCACTTAATATTAGACA 419

QY 530 TATGGTGTGTAGCCCTGCATCTCTACTCTTGTGCTTTTTCCTCCCTCCAGTGTTTGG 589
DB 420 -ATGGTGTGTAGCCCTGCATCTCTACTCTTGTGCTTTTTCCTCCCTCCAGTGTTTGG 478

QY 590 GTAGTTTGTGCTCCCTACAGCCAAAGGCAACAGATAAGTTGAGGTCTGGAGTGGCTA 649
DB 479 GTAGTTTGTGCT-CCCTACAGCCAAAGGCAACAGATAAGTTGAGGTCTGGAGTGGCTA 537

QY 650 CATAAATTTACAGCTGCAATTCCTGCTGCATCTCAGATTCACAAATGTATACAACTAAATA 709
DB 538 CATAAATTTACAGCTGCAATTCCTGCTGCATCTCAGATTCACAAATGTATACAACTAAATA 597

QY 710 CAAGTCTGCTGTTTATCACAGGAGGCTGTATCAATATAATAAATTAAGAGGGGCTG 769
DB 598 CAAGTCTGCTGTTTATCACAGGAGGCTGTATCAATATAATTAAGAGGGGCTG 657

QY 770 GTCCCATATTTGCTCTGCTGTTTGTG-----TTGTTTGTCTCTTTTCTTTTGT 820
DB 658 GT-CCATATTTGCTCTGCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTCTTTTGT 716

QY 821 GGCTCTCTCTCTCAATTTATGAAGAGAGACAGTATGTCCTCTCGGTCCTCTGA 880
DB 717 GGCTCTCTCTCTCAATTTATGAAGAGAGACAGTATGTCCTCTCGGTCCTCTGA 776

QY 881 GGGACCTGGGAGCTCAGGCTGGGAATCTCCAGGAGTAGTCCAGTCCGCTATCAAAATCAA 940
DB 940 GGGACCTGGGAGCTCAGGCTGGGAATCTCCAGGAGTAGTCCAGTCCGCTATCAAAATCAA

```

```

DB 777 GGGAGCTGGGAGCTCAGGCTGGGAATCTCCAGGAGTAGGTGCGCTATCAAAATCAA 836
QY 941 AGTCCAGGTTTCTGGGGGGAACAAACAGCCATTAACCCAGAGAGTGTGCGGCTTC 1000
DB 837 AGTCCAGGTTTCTGGGGGGAACAAACAGCCATTAACCCAGAGAGTGTGCGGCTTC 896
QY 1001 CCTCACCCAGCCTAGGCTTTTGAAGGAACAAACAGAGCAAAATGATGCGCTTC 1060
DB 897 CCTCACCCAGCCTAGGCTTTTGAAGGAACAAACAGAGCAAAATGATGCGCTTC 956
QY 1061 CTGAGGAGATTCAGCTAGAGCTCTCTCCGCCAAACCTCCCTCCGCTGAGGAAC 1120
DB 957 CTGAGGAGATTCAGCTAGAGCTCTCTCCGCCAAACCTCCCTCCGCTGAGGAAC 1016
QY 1121 TAACAAAGAAAAAAATTTGGGAAAGCAGGATTTAGAGGAAGCAAAATTCACCTGTGC 1180
DB 1017 TAACAAAGAAAAAAATTTGGGAAAGCAGGATTTAGAGGAAGCAAAATTCACCTGTGC 1076
QY 1181 CCTTGGCTCCGGAACGTGGACTAGAGAGTGTGCGGAGAGCCCGAGCCAGCCCTTC 1240
DB 1077 CCTTGGCTCCGGAACGTGGACTAGAGAGTGTGCGGAGAGCCCGAGCCAGCCCTTC 1136
QY 1241 CGCGCTCTTATAGCCGCGGCGCGGAGGAGGAGGAGCAGACCGCGGAGCCCTAA 1300
DB 1137 CGCGCTCTTATAGCCGCGGCGCGGAGGAGGAGGAGGAGCAGACCGCGGAGCCCTAA 1196
QY 1301 GACACCTGCTGTACCTCTCAC-----CCCCACCCACCCACCTCCCCCAACTCCCTAGA 1355
DB 1197 GACACCTGCTGTACCTCTCACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1256
QY 1356 TGTGCTGCGGCGCTGAACGTCGCGCTTTAAGGGGCGGCGGCGGCGGCGGCGGCGGCTTT 1415
DB 1257 TGTGCTGCGGCGCTGAACGTCGCGCTTTAAGGGGCGGCGGCGGCGGCGGCGGCGGCTTT 1316
QY 1416 CTGCTGAGTCACTCAATACATAACAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1475
DB 1317 CTGCTGAGTCACTCAATACATAACAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1376
QY 1476 ACAGGCTTGACCGATAGTAACCTCTGCGCTCGGTCAGCCGAATCTATATAAGGAA 1532
DB 1377 ACAGGCTTGACCGATAGTAACCTCTGCGCTCGGTCAGCCGAATCTATATAAGGAA 1433

RESULT 8
AF287262 201144 bp DNA linear Pal 29-APR-2001
LOCUS Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
DEFINITION and SNAP protein genes, complete cds.
ACCESSION AF287262
VERSION AF287262.1 GI:13876612
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 201144)
AUTHORS Qiu,Y., Caveller,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
TITLE Human and mouse ABCA1 comparative sequencing and transgenesis
JOURNAL studies revealing novel regulatory sequences
MEDLINE Genomics 73 (3), 66-76 (2001)
PUBMED 21251004
PUBMED 11352567
REFERENCE 2 (bases 1 to 201144)
AUTHORS Qiu,Y., Caveller,L., Chiu,S., Rubin,E. and Cheng,J.F.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-2003) Genome Science Department, Lawrence
Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
CA 94720, USA
FEATURES
source Location/Qualifiers
1..201144
/organism="Homo sapiens"
/db_xref="taxon:9606"
33931..181457
gene

```

```

mRNA
/genes="ABCA1"
join(33931..34151,58319..58476,72893..72986,77520..77661,
78931..79049,101048..101169,104152..104328,117287..117379,
123239..123299,125285..125424,125754..125870,
130079..130276,131023..131228,131749..131925,
133708..133930,135677..135898,136959..137163,
138268..138381,140179..140350,141340..141471,
142777..142919,143123..143260,143363..144183,
145441..145513,146657..146859,148527..148575,
148771..148884,150280..150428,152078..152202,
153438..153536,156568..156757,158278..158372,
159682..159714,160837..160942,162417..162491,
163013..163182,164413..164590,166588..166703,
166815..166959,168505..168628,169715..169844,
170109..170229,171017..171079,171987..172093,
174448..174589,174961..175095,176039..176142,
176625..176717,177376..177619,178560..181457)
/genes="ABCA1"
/product="ATP-binding cassette 1 sub-family A member 1"
33931..34151
/genes="ABCA1"
/number=1
58319..58476
/genes="ABCA1"
/number=2
join(58413..58476,72893..72986,77520..77661,78931..79049,
101048..101169,104152..104328,117287..117379,
123239..123299,125285..125424,125754..125870,
130079..130276,131023..131228,131749..131925,
133708..133930,135677..135898,136959..137163,
138268..138381,140179..140350,141340..141471,
142777..142919,143123..143260,143363..144183,
145441..145513,146657..146859,148527..148575,
148771..148884,150280..150428,152078..152202,
153438..153536,156568..156757,158278..158372,
159682..159714,160837..160942,162417..162491,
163013..163182,164413..164590,166588..166703,
166815..166959,168505..168628,169715..169844,
170109..170229,171017..171079,171987..172093,
174448..174589,174961..175095,176039..176142,
176625..176717,177376..177619,178560..178700)
/genes="ABCA1"
/codon_start=1
/product="ATP-binding cassette 1 sub-family A member 1"
/protein_id="AAK43526.1"
/db_xref="GI:13876613"
/translation="MACWPQLLLWNLFRRQTCOLLLEAVMLPFIPLLSVRL
SYPEQEHCFFPKAMPAGTLPWVGIIICNANPCFRPTPGAEFGVGNFKSIV
ARLFSARRLLYLQKDTSMKMRKVLRILOIKKSSNLKLDPLVDNETFSGFLYH
NLSLPKSVDMRADVILHKVFLQGYQLPLTSLCNSKSEEMOLGDOESELGCLP
KEKLAARVLRSNWDLKPLIRLTINASTPSPSKELABATKTLHSLGTLAQELFSMR
SWSDMRQVEMFLINWSSSSQIYQAVSRIVCGHPGGGLKIKSLNWEYNNKALF
GGNTEDAETFPYDNSTPYCNDLKNLLESSPLSRIWKALKP.LLVGKILYTPDIPAT
RQVAENKTFQELAVFDELEGMEELSPKIWTFMENSQEMDLVRLDSDRDHFWE
QOLGLDWTAGDIVAFIAKHDPDYQSSNGSVYTWREAFNETNOAIRTISRMEQVNLN
KLEPATEVPLINKSMELLDERKFWAGIVFTGIPGSELPHHKYKIRMDIDVNT
NKIKDWDPRADPPEDMRYVWGCPALQOVVEQALIRVLTGTEKTCVYMOOMPY
PCYVDIPFLVRSRSMPLMTLAWIYSVAIIKIGVIEKARLEIRIMLGDNSILW
F.SWFTSSLP.LLVAGLVLVILKGNLPIVSDPSVVFVLSVFAVVT.LQGLFLSTLF
SRANLAACGGIIYFTLPVLCVAMODYGFTLKIPASLSPVAFGGEGCEALFE
EQGIGVQNDLFPESVEEDGNLITVSVMILNFTFLGYMTWYIEAVPPGYGIPRPW
YEPCTKSYWGESEDEKSHPGSNOKRISEICMEEPETHLKLGVSTONLVKYVRDMKV
AVDGLALNPFEGQITSLFLGHGAGKITTMSILTGLFPPTSCYAYILGDIRESMTIR
QNLGVCPEINFLVMTVEEHLIFARLGLSKKHVKAEMEOMALDGLPSSKUSKT
SOLSCGMRKLVALFVGGSKVWILDEPTAGVDYISRGWELLKYRQRTIILT
HMDADVLGRIALISHGKLCVGGSSFLKQLGTGYLTLVKKDVSSSLSCRNSS
SVSYLVKPFDSVQSSDAGLGSDESDITLIDVSAISNLRKHVSEARLYDIGEEL
TVLVEARAKEGAFVELPHEIDRLSDIGISYGISETTLEIFILKVAEESVDAPTS
DGLTPARNRFAFGKQSLCRPPTDDAADPNDSDIDPESRETDLLSGNDGCSQYVK
GKLTQOQFVALLKRLIARSRKGFQAIVLPVAFVCIALVSLIVPPFGKYPSLE
LQPMWNYQYTVFSVSDAPDEGTGLEALLKIDPQGTFCMEGNIPDITPOAGBEEH
TTAPVPQITMDLFGQNTWMTQNPSPACQCSSDKIKMLPVPQPGAGGLPPOPKQNTA
DILQDITGRNISCLVKTVMYQIITAKSLNKKIWNVEFRYGGFSLGVSNQALPPSQEVN

```

```

exon      143963..144183
/gene="ABCA1"
/number=23
exon      145441..145513
/gene="ABCA1"
/number=24
exon      146657..146859
/gene="ABCA1"
/number=25
exon      148527..148575
/gene="ABCA1"
/number=26
exon      148771..148884
/gene="ABCA1"
/number=27
exon      150280..150428
/gene="ABCA1"
/number=28
exon      152078..152302
/gene="ABCA1"
/number=29
exon      153438..153536
/gene="ABCA1"
/number=30
exon      156568..156757
/gene="ABCA1"

```

Query Match 74.3%; Score 1138.2; DB 9; Length 201144;
Best Local Similarity 88.2%; Pred. NO. 7.6e-294;
Matches 1354; Conservative 0; Mismatches 128; Index 54; G

QY	16	GGCTGCACATGCACITCCACGGGCGCTGGTGG--CTCTTCTATGGGCTGTGCTGCAGTGTT	73
DB	32410	GGCTCCACATGCACITCCACGGGCGCTGGGGCTCTTCATCGGCTGTGCTGCAGTGTT	32469
QY	74	GATAGAACCACATCAITGTAGTACTGGGCTGTAG-CGTGGCCCTGGAGATCCTCTTCACTG	132
DB	32470	GATAGAACCACATCATGTAGTACTGGGCTGTAGCGCTGGCTGGAGATCCTGTCACTG	32529
QY	133	TAGCATGGAGGGGCGCTTGT-CASCTGAACTCTGTATGCAGGTGGTGGGAGTTCCTGGAAT	191
DB	32530	TAGCATGGAGGGGCGCTGTGCAGCTGAATCTGTGCATGCAGGTGGTGGGAGTTCCTGGAAT	32588
QY	192	ATGATGGAGCTGGAGGTGGGAAGAGAGTAGGGCTTGGGGCAGCTCTCTCATGGCACCTCA	251
DB	32590	ATGATGGAGCTGGAGGTGGGAAGAGAGTAGGGCTTGGGGCAGCTCTCTCATGGCACCTCA	32649
QY	252	TTCTGGCCAAACATCAGCTCAAACTGTGAGAGCTCTAAATGTGAACTGCGCCTTCAAGGT	311
DB	32650	TTCTGGCCAAACATCAGCTCAAACTGTGAGAGCTCTAAATGTGAACTGCGCCTTCAAGGT	32709
QY	312	GGCTACAAAGTATCTTTGTCAAGTAGGAGAGACCTTGTGGCCCTCCACGTCACTTCCAGG	371
DB	32710	GGCTACAAAGTATCTTTGTCAAGTAGGAGAGACCTTGTGGCCCTCCACGTCACTTCCAGG	32769
QY	372	GCCTGTCTT-GGCCCTTCTACGGGTCTGTCTGTAGTCTTCTATGAATCT-CCCTTCAGGG	429
DB	32770	GCCTGTCTTGGGCCCTTCTACGGGTCTGTCTGTAGTCTTCTATGAATCTGCCCTTCAGGG	32829
QY	430	CAGATTCATATTTAGACTCTTCACAGTTTGACCTTGAGTTTGGCCAGAAATAAAGTGCAT	489
DB	32830	CAGATTCATATTTAGACTCTTCACAGGGGGATCAATAGGTCGCCCATGCAIGGACTCAG	32889
QY	490	TTAGTTTGTGGCTTGATGAATGACTTAATATTTTACACATATGGTGTGTAGGCGCTGCAT	549
DB	32890	-----TNNNAGNCCCGCTCTCCCGNNNNNNNN	32916
QY	550	TCCCTACTCTTGCCTTTTCTTGGCCCTCCACAGTGTGTTTGGTAGTGTGGTCCCGCCTACA	609
DB	32917	NNNNNNNNNCCCTTTTTTTTGGCCCTCCAGTGTGTTGGGTAGTGTGCT-CCCTTACA	32975
QY	610	GCCAAAGCAACACAGATAAGTTGGAGTCTCGAGTGGCTACATAATTTTACACGACTGCA	659

[illegible]

VERSION AC012230.3 GI:7637254
KEYWORDS HTC: HTGS PHASE1: HTGS DRAFT

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
- (bases 1 to 175064)
AUTHORS
  Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE
  Homo sapiens, clone R211-1M10
JOURNAL
  Unpublished
AUTHORS
  2 (bases 1 to 175064)
  Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, K.,
  Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Boukhgalter, B.,
  Brown, A., Castle, A., Collins, S., Collins, S., Collamore, A.,
  Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
  Ferrarini, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
  Galagan, J., Gadya, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
  Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
  Lechoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
  McEwen, P., McGurk, A., McKernan, J., McLaughlin, J., MeGrath, J.,
  Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
  Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
  Stange-Thoman, N., Stojanovic, N., Subramanian, A., Talamas, J.,
  Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
  Wyman, D., Ye, W. J., Zimmer, A. and Zody, N.
Direct Submission
Submitted (21-Oct-1999), Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
  Smith, A.P.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2510
Center clone name: 1.M.10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 145000; agarose-fp
Insert size: 171264; sum-of-contigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1003: contig of 1003 bp in length
* 1004 1103: gap of 100 bp
* 1104 2634: contig of 1531 bp in length
* 2635 2734: gap of 100 bp
* 2735 4415: contig of 1681 bp in length
* 4416 4515: gap of 100 bp
* 4516 5785: contig of 1270 bp in length
* 5786 5885: gap of 100 bp
* 5886 7879: contig of 1994 bp in length
* 7880 9797: gap of 100 bp
* 9780 9685: contig of 1707 bp in length
* 9687 9786: gap of 100 bp
* 9787 12253: contig of 2467 bp in length
* 12254 12353: gap of 100 bp
* 12354 15228: contig of 2875 bp in length
* 15229 15328: gap of 100 bp
* 15329 17200: contig of 1872 bp in length

```

```

* 17201 17300: gap of 100 bp
* 17301 20131: contig of 2831 bp in length
* 20132 20231: gap of 100 bp
* 20232 22587: contig of 2356 bp in length
* 22588 22687: gap of 100 bp
* 22688 25707: contig of 3020 bp in length
* 25708 25807: gap of 100 bp
* 25808 28184: contig of 2377 bp in length
* 28185 28284: gap of 100 bp
* 28285 31338: contig of 3054 bp in length
* 31339 31438: gap of 100 bp
* 31439 34299: contig of 2861 bp in length
* 34300 34398: gap of 100 bp
* 34400 36318: contig of 3919 bp in length
* 36319 38418: gap of 100 bp
* 38419 42835: contig of 4417 bp in length
* 42836 42935: gap of 100 bp
* 42936 45448: contig of 2513 bp in length
* 45449 45548: gap of 100 bp
* 45549 48116: contig of 2568 bp in length
* 48117 48216: gap of 100 bp
* 48217 52618: contig of 4402 bp in length
* 52619 52718: gap of 100 bp
* 52719 56592: contig of 3874 bp in length
* 56593 56692: gap of 100 bp
* 56693 59635: contig of 2943 bp in length
* 59636 59735: gap of 100 bp
* 59736 59661: contig of 3926 bp in length
* 63662 63761: gap of 100 bp
* 63762 68437: contig of 4576 bp in length
* 68438 68537: gap of 100 bp
* 68538 71458: contig of 2921 bp in length
* 71459 71558: gap of 100 bp
* 71559 76888: contig of 5330 bp in length
* 76889 76988: gap of 100 bp
* 76989 82113: contig of 5125 bp in length
* 82114 82213: gap of 100 bp
* 82214 88220: contig of 6007 bp in length
* 88221 88320: gap of 100 bp
* 88321 93493: contig of 5179 bp in length
* 93500 93599: gap of 100 bp
* 97901 97901: contig of 4302 bp in length
* 97902 98001: gap of 100 bp
* 98002 103016: contig of 5015 bp in length
* 103017 103116: gap of 100 bp
* 103117 109178: contig of 6062 bp in length
* 109179 109278: gap of 100 bp
* 109279 117307: contig of 8029 bp in length
* 117308 117407: gap of 100 bp
* 117408 124079: contig of 6672 bp in length
* 124080 124179: gap of 100 bp
* 124180 131281: contig of 7102 bp in length
* 131282 131381: gap of 100 bp
* 131382 138059: contig of 6578 bp in length
* 138060 138159: gap of 100 bp
* 138160 145491: contig of 7332 bp in length
* 145492 145591: gap of 100 bp
* 145592 157391: contig of 11800 bp in length
* 157392 157491: gap of 100 bp
* 157492 175064: contig of 17573 bp in length.
FEATURES
  Location/Qualifiers
    1..175064
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="R211-1M10"
      /clone_lib="RFCI-11 Human Male BAC"
    misc_feature
      1..1003
        /note="assembly_fragment"
    misc_feature
      1104..2634
        /note="assembly_fragment"
    misc_feature
      2735..4415
        /note="assembly_fragment"
    misc_feature
      4516..5785

```

[illegible]

Db	3034	GCCTATCAAAATCAAGTCCAGTTTGCTGGGGGGAACAAAGACAGCCCATACCAG	3093
QY	985	AGGACTTCGGCTTCCTCCCTCACCCAGAGCTAGCCCTTTGAAGGAAACAAAGACACAGA	1044
Db	3094	AGGACTTCGGCTTCCTCCCTCACCCAGAGCTAGCCCTTTGAAGGAAACAAAGACACAGA	3153
QY	1045	CAAAATGATTGGCTCTCTGAGGAGATTCAGCCTAGAGCTCTCTCCGCCCAATCCCTCC	1104
Db	3154	CAAAATGATTGGCTCTCTGAGGAGATTCAGCCTAGAGCTCTCTCCGCCCAATCCCTCC	3213
QY	1105	CTCGGCTGAGGAACATAACAAAGGAAAAAATTCGGGAAAGCAGGATTAGAGGAAG	1164
Db	3214	CTCGGCTGAGGAACATAACAAAGGAAAAAATTCGGGAAAGCAGGATTAGAGGAAG	3273
QY	1165	CAAAATGATTGGCTCTCTGAGGAGATTCAGCCTAGAGCTCTCTCCGCCCAATCCCTCC	1224
Db	3274	CAAAATGATTGGCTCTCTGAGGAGATTCAGCCTAGAGCTCTCTCCGCCCAATCCCTCC	3333
QY	1225	CGAGCCAGAGCTTCCTCGCGCTTTAGCCGCGGGCCCGGGCGGGGGAAGGGAGCC	1284
Db	3334	CGAGCCAGAGCTTCCTCGCGCTTTAGCCGCGGGCCCGGGCGGGGGAAGGGAGCC	3393
QY	1285	AGACCGCGGACCTTAAGACACCTGCTGACCTCCACCCCAACCCACCCACCTCCCCCC	1344
Db	3394	AGACCGCGGACCTTAAGACACCTGCTGACCTCCACCCCAACCCACCCACCTCCCCCC	3453
QY	1345	AACTCCCTAGATGCTGCTGGCGGCTGAACTGTCGCGCGTTAAGGGGCGGGCCCGGCT	1404
Db	3454	AACTCCCTAGATGCTGCTGGCGGCTGAACTGTCGCGCGTTAAGGGGCGGGCCCGGCT	3513
QY	1405	CCACCTGCTTCTGCTGAGTACTGAATACATAACAGAGCCGGGACGCGGGGGA	1464
Db	3514	CCACCTGCTTCTGCTGAGTACTGAATACATAACAGAGCCGGGACGCGGGGGA	3573
QY	1465	GGAGGAGAGCAGAGCTTTGACGATAGTACCTTCGCGTTCGCGAGCGGAATCTAT	1524
Db	3574	GGAGGAGAGCAGAGCTTTGACGATAGTACCTTCGCGTTCGCGAGCGGAATCTAT	3633
QY	1525	AAAGGAA 1532	
Db	3634	AAAGGAA 3641	
RESULT 10			
HSAS252201		1167 bp DNA linear	PRI 10-APR-2001
LOCUS			
DEFINITION			
ACCESSTION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
SOURCE			


```

/db_xref="taxon:9606"
/chromosome="9"
/map="9q22-31"
/cell_type="leukocyte"
1..1167
/gene="ABC-1"
promoter
1..942
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
896..900
/gene="ABC-1"
1148..1167
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
278 a 315 c 327 g 247 t

Query Match 58.4%; Score 893.4; DB 9; Length 1167;
Best Local Similarity 99.8%; Pred. No. 6.6e-229;
Matches 906; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 626 TAAGTGGAGGCTGGAGTGGCTACATAAATTTACAGAGCTGCAATTTCTGGCTGCACI 685
DB 1 TAAGTGGAGGCTGGAGTGGCTACATAAATTTACAGAGCTGCAATTTCTGGCTGCACI 60
QY 686 TCACAAATGTATACAACTAAACAAAGTCTGCTGTTTATACAGAGGAGCTGATCAA 745
DB 61 TCACAAATGTATACAACTAAACAAAGTCTGCTGTTTATACAGAGGAGCTGATCAA 220
QY 746 TATAATGAATTAAGGGGCTGCTCCATATATTTGCTGTTTGTGTTTGTGTTTTC 805
DB 121 TATAATGAATTAAGGGGCTGCTCCATATATTTGCTGTTTGTGTTTGTGTTTTC 179
QY 806 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTTC 865
DB 180 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTTC 239
QY 866 TCTGGGTCCTCTGAGGAGCTGGGAGCTCAGGTCGGAATTTCAAGAGAGCAAGTGTTC 925
DB 240 TCTGGGTCCTCTGAGGAGCTGGGAGCTCAGGTCGGAATTTCAAGAGAGCAAGTGTTC 299
QY 926 CCTATCAAAATCAAGTCCAGGTTTGTGGGGGAAACAAAGCAGCCATACCCAGA 985
DB 300 CCTATCAAAATCAAGTCCAGGTTTGTGGGGGAAACAAAGCAGCCATACCCAGA 359
QY 986 GCATGTCGCTTCCCTCAGCCGAGCTTGGGCTTTGAAAGGAAACAAAGCAAGAC 1045
DB 360 GCATGTCGCTTCCCTCAGCCGAGCTTGGGCTTTGAAAGGAAACAAAGCAAGAC 429
QY 1046 AAAATGATTTGGGCTCTGAGGAGATTCAGCTTCTCTCTCCCAATCCCTCC 1105
DB 420 AAAATGATTTGGGCTCTGAGGAGATTCAGCTTCTCTCTCCCAATCCCTCC 479
QY 1106 TCCGGCTAGGAACTAACAAAGGAAACAAATTCGGAAGCAGATTTAGAGGAGC 1165
DB 480 TCCGGCTAGGAACTAACAAAGGAAACAAATTCGGAAGCAGATTTAGAGGAGC 539
QY 1166 AATTTCAGTGTGCTTGGCTGCGGAACTGAGAGTCTGCGGCGCAGGCC 1225
DB 540 AATTTCAGTGTGCTTGGCTGCGGAACTGAGAGTCTGCGGCGCAGGCC 599
QY 1226 CGAGCCAGGCTTCCCGGCGCTTAGGCGCGGCGCGCGCGGCGGCGGAGGAGCGCA 1285
DB 600 CGAGCCAGGCTTCCCGGCGCTTAGGCGCGGCGCGCGCGGCGGCGGAGGAGCGCA 659
QY 1286 GACCGGAGCCCTTAAGACACTGCTGTACCTCCACCCCGCAGCCACCTCCGCCCA 1345
DB 660 GACCGGAGCCCTTAAGACACTGCTGTACCTCCACCCCGCAGCCACCTCCGCCCA 719
QY 1346 ACTCCCTAGATGTGCTGGGCGCTCAAGCTGCGCGCTTAAAGGGCGGCGCGCGCTC 1405
DB 720 ACTCCCTAGATGTGCTGGGCGCTCAAGCTGCGCGCTTAAAGGGCGGCGCGCGCTC 779

```

```

QY 1406 CACGTGCTTTCTGCTGAGTGACTGAATACATAAAGAGGCGCGGCGGGGAG 1465
DB 780 CACGTGCTTTCTGCTGAGTGACTGAATACATAAAGAGGCGCGGCGGGGAG 839
QY 1466 GAGGGAGACAGAGCTTTGACGATAGTAACCTCTGGCTCGGTGACGCGGAATCTA 1525
DB 840 GAGGGAGACAGAGCTTTGACGATAGTAACCTCTGGCTCGGTGACGCGGAATCTA 899
QY 1526 AAAGGAA 1532
DB 900 AAAGGAA 906

RESULT 11
AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
LOCUS Homo sapiens ATP3 binding cassette transporter 1 (ABCA1) gene,
DEFINITION Promoter and exon 1.
ACCESSION AF258623
VERSION AF258623.2 GI:8677405
KEYWORDS
SEGMENT
SOURCE i of 4
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 224 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REFERENCE 3 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
REMARK Sequence update by submitter
COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.
FEATURES
Location/Qualifiers
source
1..1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
224..844
/gene="ABCA1"
845..1147
/gene="ABCA1"
/number=1
BASE COUNT 278 a 313 c 328 g 248 t
ORIGIN
Query Match 58.3%; Score 893.4; DB 9; Length 1167;
Best Local Similarity 99.8%; Pred. No. 1.8e-228;
Matches 905; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 626 TAAGTGGAGGCTGGAGTGGCTACATAAATTTACAGAGCTGCAATTTCTGGCTGCACI 685
DB 1 TAAGTGGAGGCTGGAGTGGCTACATAAATTTACAGAGCTGCAATTTCTGGCTGCACI 60
QY 686 TCACAAATGTATACAACTAAATACAAAGTCTGCTGTTTATACAGAGGAGCTGATCAA 745
DB 61 TCACAAATGTATACAACTAAATACAAAGTCTGCTGTTTATACAGAGGAGCTGATCAA 120
QY 746 TATAATGAATTAAGGGGCTGCTCCATATATTTGCTGTTTGTGTTTGTGTTTTC 805
DB 121 TATAATGAATTAAGGGGCTGCTCCATATATTTGCTGTTTGTGTTTGTGTTTTC 179
QY 806 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTTC 865
DB 180 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTTC 239
QY 866 TCTGGGTCCTCTGAGGAGCTGGGAGCTCAGGTCGGAATTTCAAGAGAGCAAGTGTTC 925
DB 240 TCTGGGTCCTCTGAGGAGCTGGGAGCTCAGGTCGGAATTTCAAGAGAGCAAGTGTTC 299
QY 926 CCTATCAAAATCAAGTCCAGGTTTGTGGGGGAAACAAAGCAGCCATACCCAGA 985
DB 300 CCTATCAAAATCAAGTCCAGGTTTGTGGGGGAAACAAAGCAGCCATACCCAGA 359
QY 986 GCATGTCGCTTCCCTCAGCCGAGCTTGGGCTTTGAAAGGAAACAAAGCAAGAC 1045
DB 360 GCATGTCGCTTCCCTCAGCCGAGCTTGGGCTTTGAAAGGAAACAAAGCAAGAC 429
QY 1046 AAAATGATTTGGGCTCTGAGGAGATTCAGCTTCTCTCTCCCAATCCCTCC 1105
DB 420 AAAATGATTTGGGCTCTGAGGAGATTCAGCTTCTCTCTCCCAATCCCTCC 479
QY 1106 TCCGGCTAGGAACTAACAAAGGAAACAAATTCGGAAGCAGATTTAGAGGAGC 1165
DB 480 TCCGGCTAGGAACTAACAAAGGAAACAAATTCGGAAGCAGATTTAGAGGAGC 539
QY 1166 AATTTCAGTGTGCTTGGCTGCGGAACTGAGAGTCTGCGGCGCAGGCC 1225
DB 540 AATTTCAGTGTGCTTGGCTGCGGAACTGAGAGTCTGCGGCGCAGGCC 599
QY 1226 CGAGCCAGGCTTCCCGGCGCTTAGGCGCGGCGCGCGCGGCGGAGGAGCGCA 1285
DB 600 CGAGCCAGGCTTCCCGGCGCTTAGGCGCGGCGCGCGCGGCGGAGGAGCGCA 659
QY 1286 GACCGGAGCCCTTAAGACACTGCTGTACCTCCACCCCGCAGCCACCTCCGCCCA 1345
DB 660 GACCGGAGCCCTTAAGACACTGCTGTACCTCCACCCCGCAGCCACCTCCGCCCA 719
QY 1346 ACTCCCTAGATGTGCTGGGCGCTCAAGCTGCGCGCTTAAAGGGCGGCGCGCGCTC 1405
DB 720 ACTCCCTAGATGTGCTGGGCGCTCAAGCTGCGCGCTTAAAGGGCGGCGCGCGCTC 779

```

```

2 (bases 1 to 69570)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Balwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boughgatter,B., Brown,A., Burkett,G., Castle,A.,
Chapel,V., Colangelo,C., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,K., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,T., Lehotzky,J., Levine,K., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,I., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,I., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severi,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talmans,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,R., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:670587.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WISR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25:2
Center clone name: L_NJ0
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
1 871: contig of 871 bp in length
972 971: gap of 100 bp
972 1834: contig of 863 bp in length
1835 1934: gap of 100 bp
1935 2804: contig of 870 bp in length
2805 2904: gap of 100 bp
2905 3745: contig of 841 bp in length
3746 3845: gap of 100 bp
3846 4696: contig of 851 bp in length
4697 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 6540: contig of 800 bp in length:
6541 6640: gap of 100 bp
6641 7509: contig of 869 bp in length
7510 7609: gap of 100 bp
7610 8479: contig of 870 bp in length
8480 8579: gap of 100 bp
8580 9430: contig of 851 bp in length
9431 9530: gap of 100 bp
9531 10376: contig of 846 bp in length
10377 10476: gap of 100 bp
10477 11322: contig of 846 bp in length
11323 11422: gap of 100 bp
11423 12302: contig of 880 bp in length
12303 12402: gap of 100 bp
12403 13280: contig of 878 bp in length
13281 13380: gap of 100 bp
*

```

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
745	TATAATGAATTAAGAGGGGCTGGTCCCATATTTGCTGTGTTTGTGTTTGTGTTTTC	805
121	TATAATGAATTAAGAGGGGCTGGT-CATATTTGCTGTGTTTGTGTTTGTGTTTTC	179
806	TTTTTTGTTTGTGTTGTCCTTCCCTCAATTAAGAGAGAGAGAGAGAGAGAGAG	865
180	TTTTTTGTTTGTGTTGTCCTTCCCTCAATTAAGAGAGAGAGAGAGAGAGAGAG	239
866	TCCTGGGTCCTCTGAGGAGACCTGGGAGCTCAGGCTGGGAATCTCCAAAGCAGTAGGTG	925
240	TCCTGGGTCCTCTGAGGAGACCTGGGAGCTCAGGCTGGGAATCTCCAAAGCAGTAGGTG	299
926	CCATCAAAATCAAAAGTCCAGTTTGTGGGGGAGAAACAAAAGCAGCCCATACCCAGA	985
300	CCATCAAAATCAAAAGTCCAGTTTGTGGGGGAGAAACAAAAGCAGCCCATACCCAGA	359
986	GGAGTCTCCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	1045
360	GGAGTCTCCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	419
1046	AAATGATTTGGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	1105
420	AAATGATTTGGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	479
1106	TCCGGCTGAGGAAACTAACAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1165
480	TCCGGCTGAGGAAACTAACAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	539
1166	AAATTCACACTGGTGGCTTGGCTGGGAGAAAGTGGAGTGGAGTGGAGTGGAGTGG	1225
540	AAATTCACACTGGTGGCTTGGCTGGGAGAAAGTGGAGTGGAGTGGAGTGGAGTGG	599
1226	CGAGCCAGAGCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	1285
600	CGAGCCAGAGCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	659
1286	GACCGCGAGCCTTAAGACACCTGTGTACCTTCCACCTTCCACCTTCCACCTTCC	1345
660	GACCGCGAGCCTTAAGACACCTGTGTACCTTCCACCTTCCACCTTCCACCTTCC	719
1346	ACTCCCTAGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	1405
720	ACTCCCTAGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	779
1406	CACGTGCTTTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	1465
780	CACGTGCTTTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	839
1466	GAGGAGAGACAGGCTTTGACGATAGTAACCTCTGCGCTCGGTCGAGCCGAAATCTATA	1525
840	GAGGAGAGACAGGCTTTGACGATAGTAACCTCTGCGCTCGGTCGAGCCGAAATCTATA	999
1526	AAAGGAA 1532	
900	AAAGGAA 906	
RESULT 12	ACC2:246	69570 bp DNA linear HTG 13-JUL-2000
LOCUS	HOMO sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.	
DEFINITION	AC021246	
ACCESSION	AC021246	
VERSION	AC021246.2 GI:9119882	
KEYWORDS	HTG; HTGS_PHASE0.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	I (bases 1 to 69570)	
JOURNAL	Siren, P., Linton, L., Nusbaum, C. and Lander, E.	
	HOMO sapiens chromosome, clone RP11-1N10	
	Unpublished	

```
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19968: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21865: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 25621: contig of 851 bp in length
* 26222 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 45724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 46643: contig of 819 bp in length
* 46644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47699: gap of 100 bp
* 47700 48551: contig of 852 bp in length
```

```
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55328: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67551: contig of 867 bp in length
* 67552 67751: gap of 100 bp

Query Match      26.7%; Score 408.8; DB 2; Length 69570;
Best Local Similarity 99.1%; Pred. No. 3.5e-98;
Matches 432; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1099 CCTTCCCTCGGCTGAGAACTAACAAA-GGAAAAAAAATTCGGAAAGCAGGATTTA 1157
DB 41109 CCTTCCCTCGGCTGAGAACTAACAAAAGGGAAGGAAAAATTCGGAAAGCAGGATTTA 41168

QY 1158 GAGGAAGCAAAATTCCTACTGGTGCCTTGGCTGCCGGGACGCTGGACTAGAGCTTCGGG 1217
DB 41169 GAGGAAGCAAAATTCCTACTGGTGCCTTGGCTGCCGGGACGCTGGACTAGAGCTTCGGG 41228

QY 1218 CGAGGCCCGAGCCCGAGCGCTTCCCGCGCTCTTAGCCGCGGCGCGCGGGGGGAAG 1277
DB 41229 CGAGGCCCGAGCCCGAGCGCTTCCCGCGCTCTTAGCCGCGGCGCGCGGGGGGAAG 41288

QY 1278 GGGAGCGAGACCGCGGACCGCTTAAGACACCTGCTGATCCCTCCACCCCGCCACCCCAACC 1337
DB 41289 GGGAGCGAGACCGCGGACCGCTTAAGACACCTGCTGATCCCTCCACCCCGCCACCCCAACC 41348

QY 1338 TCCCCCAACT-CCCTAGATGTGTCGTGGCGGCTGAACGTCGCCCGCTTTAAGGGCGGG 1396
DB 41349 TCCCCCAACTCCCTATATGTCGTGGCGGCTGAACGTCGCCCGCTTTAAGGGCGGG 41408

QY 1397 CCGCGGCTCCAGCTGCTTCTGCTGAGTGACTGAATACATACAGACGCCCGGGACGG 1456
DB 41409 CCGCGGCTCCAGCTGCTTCTGCTGAGTGACTGAATACATACAGACGCCCGGGACGG 41468

QY 1457 GCGGGGAGGAGGAGGAGGACAGGCTTTGACCGGATAGTAACCTCTGCGCTCGGTGCAGCC 1516
DB 41469 GCGGGGAGGAGGAGGAGGACAGGCTTTGACCGGATAGTAACCTCTGCGCTCGGTGCAGCC 41528
```

```

QY 1517 GAATCTATAAAGGAA 1532
      |.....|
DB 41529 GAATCTATAAAGGAA 41544

RESULT 13
AC021246/c
LOCUS
DEFINITION Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021246
VERSION 2
KEYWORDS HTG; HTGS-PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69570)
Birten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-1N10
Unpublished
2 (bases 1 to 69570)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Becker,I., Beda,F.,
Boguslavsky,L., Bouknight,B., Brown,A., Burkett,G., Castle,A.,
Chapel,Y., Coangel,G., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lebecky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McManus,P., McGurk,A., McKernan,K.,
McPheeters,R., Melidrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence-submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: 1_N10
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 871: contig of 871 bp in length
* 872 971: gap of 100 bp
* 972 1834: contig of 863 bp in length
* 1835 1934: gap of 100 bp
* 1935 2804: contig of 870 bp in length
* 2805 2904: gap of 100 bp
* 2905 3745: contig of 841 bp in length

```


* 1011 1873: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length
* 2825 2924: gap of 100 bp
* 2923 3802: contig of 878 bp in length
* 3803 3902: gap of 100 bp
* 3903 4816: contig of 914 bp in length
* 4817 4916: gap of 100 bp
* 4917 5759: contig of 843 bp in length
* 5760 5859: gap of 100 bp
* 5860 6764: contig of 905 bp in length
* 6765 6864: gap of 100 bp
* 6865 7747: contig of 883 bp in length
* 7748 7847: gap of 100 bp
* 7848 8755: contig of 908 bp in length
* 8756 8855: gap of 100 bp
* 8856 9753: contig of 898 bp in length
* 9754 9853: gap of 100 bp
* 9854 10757: contig of 904 bp in length
* 10758 10857: gap of 100 bp
* 10858 11732: contig of 875 bp in length
* 11733 11832: gap of 100 bp
* 11833 12738: contig of 907 bp in length
* 12740 12839: gap of 100 bp
* 12840 13710: contig of 871 bp in length
* 13711 14684: contig of 874 bp in length
* 14685 14784: gap of 100 bp
* 14785 15662: contig of 878 bp in length
* 15663 15762: gap of 100 bp
* 15763 16677: contig of 915 bp in length
* 16678 16777: gap of 100 bp
* 16778 17678: contig of 901 bp in length
* 17679 17778: gap of 100 bp
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18780 19632: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19733 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22573: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23688: gap of 100 bp
* 23689 24554: contig of 886 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length
* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length

* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
* 54226 54325: gap of 100 bp
* 54326 55208: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match 8.9%: Score 135.6; DB 2; Length 90598;
Best Local Similarity 55.1%: Pred. No. 7.5e-25;
Matches 172: Conservative 0; Mismatches 138; Indels 2; Gaps 2;

[illegible]

```

fragment_chain:2"
58385..60786
/note="assembly_fragment:01660
fragment_chain:2"
60887..66707
/note="assembly_fragment:00374
fragment_chain:2"
66808..92102
/note="assembly_fragment:01792
fragment_chain:2"
92203..99909
/note="assembly_fragment:01082
fragment_chain:2"
100010..103760
/note="assembly_fragment:00358
fragment_chain:2"
103861..110765
/note="assembly_fragment:01182
fragment_chain:3"
110866..114685
/note="assembly_fragment:01582
fragment_chain:3"
114786..125959
/note="assembly_fragment:00043
fragment_chain:3"
126060..128749
/note="assembly_fragment:00352
fragment_chain:3"
128850..132373
/note="assembly_fragment:00217
fragment_chain:4"
132474..134930
/note="assembly_fragment:01737
fragment_chain:4"
135031..139301
/note="assembly_fragment:00487"
139402..144448
/note="assembly_fragment:00557"
144549..147074
/note="assembly_fragment:01342"
147175..156317
/note="assembly_fragment:01564.0"
156418..186889
/note="assembly_fragment:01773"
COUNT 52900 a 39593 c 40710 g 51338 t 2348 others

```

Query Match	8.4%;	Score 128.8;	DB 2;	Length 186889;
Best Local Similarity	79.0%;	Pred. NC. 5.6e-23;		
Matches 166;	Conservative 0;	Mismatches 42;	Indels 2;	Gaps 1;
QY 1325	CACCCACCCACCTCCGCCCACTGCTAGATGTGCTGGGGGCTGAACGTGCCCGT	1384		
Db 99329	CCCCACACAGCTCTACACATCCGCCCACTGCCCGCTGGCGGCACACTGTGCCGCT	99270		
QY 1385	TTAAGGGGGGGGGGGGGGCTCCACGTCCTTCTGCTCAGTCACTGACTACATAACAGA	1444		
Db 99269	TTAAGGGGGGGGGCATGCTCCACGCTTCTCTCAGTGACTGAACACTACATAACAGA	99210		
QY 1445	GGCGGGGAACGGGGCGGG--AGGAGGAGAGCACAGCTTGACCGCATACTACCTCTG	1502		
Db 99209	GGCGGGGAAGGGGGGGGGAAGGAGGAGAGACGCGTTGACCGGTACTACCCCGG	99150		
QY 1503	CGCTCGGTGCAGCGGAATCTATAAAGGAA	1532		
Db 99149	CGTCCGACAGCGGAATCTATAAAGGAA	99120		

Search completed: February 4, 2003, 01:14:38
Job time : 5324.17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:46:48 ; Search time 126.452 Seconds
(without alignments)
8245.650 Million cell updates/sec

Title: US-09-596-141C-3_COPY_1181_1643
Perfect score: 463
Sequence: 1 ccttgctccgggaacgtg.....cgggcaggcggggagcctc 463

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues 4370478

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101032.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	1197	24	ABL58400 Human large ATP-binding cassette transporter 1(hABCL) promoter sequence.
2	463	100.0	1643	22	AAH07432 Human ABC1 gene ex
3	463	100.0	1643	22	AAH18506 Human cDNA clone (
4	447	96.5	18399	22	AAH04729 Human cDNA clone (
5	446.4	96.4	3231	24	AAH17451 Human cDNA sequenc
6	372.4	80.4	29.0	24	AAH57265 Human ABC1 transcr
7	119.4	25.8	227	21	AAH37267 Human ABC1 transcr
8	98	21.2	7260	22	AAH09615 Human secreted pro
9	93	21.2	7260	22	AAH21326 Human ATP binding
					Human ATP binding

ALIGNMENTS

RESULT 1
ABL58400
ID ABL58400 standard; DNA; 1197 BP.
XX
AC ABL58400;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human large ATP-binding cassette transporter 1(hABCL) promoter sequence.
XX
KW Human: large ATP-binding cassette transporter 1; ABC1; promoter;
XX antiarteriosclerotic; gene transfer; transactivator; ds.
XX
OS Homo sapiens.
XX
PN WO200183505-A1
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13654.
XX
PR 28-APR-2000; 2000US-0560372.
XX
XX (UYCO) UNIV COLLEGE NEW YORK.
XX
PA Nucleotide sequenc
XX
PI Tail AR;
XX
DR WPI; 2002-049334/06.
XX
XX Novel isolated human large ATP-binding cassette transporter 1 promoter
XX capable of directing transcription of heterologous coding sequence
XX positioned downstream to it, useful for expressing foreign DNA in host
PT

```
cells -
XX
PS Claim 1; Fig 3; 68pp; English.
XX
CC The invention relates to an isolated human large ATP-binding cassette
CC transporter 1 (ABCI) promoter capable of directing transcription of
CC heterologous coding sequence positioned downstream to it. The ABCI
CC promoter is useful for expressing foreign DNA in a host cell, by
CC introducing into the host cell a gene transfer vector comprising the
CC promoter operably linked to a foreign DNA encoding a desired polypeptide
CC or RNA, where the foreign DNA is expressed. The gene transfer can be
CC introduced into the host cell by adenovirus infection, liposome-mediated
CC transfer, topical application to the cell or microinjection. The gene
CC transfer vector encodes and expresses a reporter molecule. The method
CC further involves introducing into the cell a gene transfer vector
CC comprising a nucleic acid segment encoding a transactivator protein
CC capable of upregulating the ABCI promoter, or contacting the cell with
CC the transactivator protein, or an agonist of the transactivator protein.
CC Modulators of human ABCI gene expression are useful for treating
CC atherosclerosis. The present sequence represents the ABCI promoter.
XX
SQ Sequence 1197 BP; 284 A; 314 C; 328 G; 271 T; 0 other;
Query Match 100.0%; Score 463; DB 24; Length 1197;
Best Local Similarity 100.0%; Pred. No. 7.8e-108;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTTGGCTGCCGGGAACGTGGACTAGAGAGTCTGCGGCGACGCCCGAGCCGAGCGCTTC 60
Df 719 CCTTGGCTGCCGGGAACGTGGACTAGAGAGTCTGCGGCGACGCCCGAGCCGAGCGCTTC 778
QY 61 CGCGCGGTCTTAGCGCGCGGCGCGCGCGGCGGGAAGGAGGACGACGCGCGGACCGCTAA 120
Df 779 CGCGCGGTCTTAGCGCGCGGCGCGCGGCGGGAAGGAGGACGACGCGCGGACCGCTAA 838
QY 121 GACACCTGCTGTACCTCCACCCCGACCCACCCACCCACCTCCCGCCACCTCCCTAGATGTG 180
Df 839 GACACCTGCTGTACCTCCACCCCGACCCACCCACCCACCTCCCGCCACCTCCCTAGATGTG 898
QY 181 CGTGGCGGCTGAACCTGCGCGCTTTAAGGCGGCGCGCGCGCGCTCCACCTGCTTCGTCT 240
Df 899 CGTGGCGGCTGAACCTGCGCGCTTTAAGGCGGCGCGCGCGCGCTCCACCTGCTTCGTCT 958
QY 241 GAGTGAATGACTACATAACAGAGCGCGGACGCGGCGGCGGAGGAGGAGGACAGCAGG 300
Df 959 GAGTGAATGACTACATAACAGAGCGCGGACGCGGCGGCGGAGGAGGAGGAGGACAGCAGG 1018
QY 301 CTTTGACCGATAGTAACCTCTGCGCTCGGCGAGCGCGGCGGCGGCGGCGGAGGAGGAGCAGG 1078
Df 1019 CTTTGACCGATAGTAACCTCTGCGCTCGGCGAGCGCGGCGGCGGCGGCGGAGGAGGAGCAGG 1138
QY 361 GGCATAAAGCCCGTAATTCGAGCGAGAGTGTAGTGGGCGGCGGCGGCGGAGCGGAGCC 420
Df 1079 GGCATAAAGCCCGTAATTCGAGCGAGAGTGTAGTGGGCGGCGGCGGCGGAGCGGAGCC 1138
QY 421 GACCCCTCTCTCCCGGCTGCGGCGAGGCGGCGGCGGCGGAGCTC 463
Df 1139 GACCCCTCTCTCCCGGCTGCGGCGAGGCGGCGGCGGAGCTC 1181
RESULT 2
AAF24681
ID AAF24681 standard; DNA; 1643 BP.
XX
AC AAF24681;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the 5' flanking region of the human ABCI gene.
KW Human; adenosine triphosphate binding cassette protein 1; ABCI;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
```

QY 361 GCACAAACCCCGTAATTGCGAGCGAGAGTGGTGGGGCGGAGCCCGGAGCGGAGCC 420
|||||
Db 1541 GSCAAAACCCCGTAATTGCGAGCGAGAGTGGTGGGGCGGAGCCCGGAGCGGAGCC 1600
|||||
QY 421 GACCCCTTCTCTCCCGGCTGCGGCGAGGCGAGGGCGGGGAGCTC 463
|||||
Db 1601 GACCCCTTCTCTCCCGGCTGCGGCGAGGCGAGGGCGGGGAGCTC 1643
|||||

RESULT 3

AAE24703
ID AAE24703 standard; DNA; 1643 BP.
XX AC
XX AAE24703;
DI 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the 5' flanking region of the human ABC1 gene.
XX
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; SS.
XX
OS Homo sapiens.
XX
PN WO200078971-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16591.
XX
PR 18-JUN-1999; 99US-0140264.
PR 14-SEP-1999; 99US-0153872.
PR 19-NOV-1999; 99US-0166573.
XX
PA (CVTH-) CV THERAPEUTICS INC.
PA (UNIW) UNIV WASHINGTON.
XX
PI Lawn RM, Wade D, Oram JF, Garvin M;
XX
DI WPI; 2001-13781/14.
XX

Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
polynucleotides and polypeptides, useful for treatment of heart disease
and other disorders associated with hypercholesterolemia and
atherosclerosis -

Disclosure; page 138-139; 211pp; English.

The present sequence represents the 5' flanking region of the human
adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1
resides in cell membranes and utilizes ATP hydrolysis to transport a wide
variety of substrates across the plasma membrane. ABC1 is a pivotal
protein in the apolipoprotein-mediated mobilisation of intracellular
cholesterol stores. ABC1 is defective in Tangier disease, a genetic
disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1
gene is localised to chromosome 9q22-q31. The ABC1 genes and proteins
are useful for developing pharmaceutical agents for the treatment of
heart disease and other disorders associated with hypercholesterolemia
and atherosclerosis. The genes are useful for developing screening assays
to screen for compounds that regulate the expression of genes associated
with cholesterol transport. The genes and proteins are also useful for
CC are also useful as diagnostic indicators of cardiovascular disease and
other disorders associated with hypercholesterolemia.

Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match:
Best Local Similarity 100.0%; Score 463; DB 22; Length 1643;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTGGCTCCCGGACGTTGGACTAGAGACTCTGGCGGCGACCCCGGAGCCCGGCTTC 60

Db 1181 CTTTGGCTCCCGGACGTTGGACTAGAGACTCTGGCGGCGACCCCGGAGCCCGCTTC 1240
|||||
QY 61 CGCGCGCTTTAGCGCGCGGCGCGCGGGGGAAGGGAGCGAGACCGCGGACCTAA 120
|||||
Db 1241 CGCGCGCTTTAGCGCGCGGCGCGCGGGGGAAGGGAGCGAGACCGCGGACCTAA 1300
|||||
QY 121 GACACCTGCTGTACCTCCACCCCGACCCACCCACCTCCCGGAGCTCCCTAGATGTCT 180
|||||
Db 1301 GACACCTGCTGTACCTCCACCCCGACCCACCTCCCGGAGCTCCCTAGATGTCT 1360
|||||
QY 181 CGTGGCGGCTGAACGTCGCGCGCTTTAAGGGCGGGCGCGCGGCTCCACGTGTTTCTGCT 240
|||||
Db 1361 CGTGGCGGCTGAACGTCGCGCGCTTTAAGGGCGGGCGCGCGGCTCCACGTGTTTCTGCT 1420
|||||
QY 241 GAGTGACTGAACCTAGATAAAGAGAGCGCGGGAACCGGGCGGGGAGGAGGAGACACAGG 300
|||||
Db 1421 GAGTGACTGAACCTAGATAAAGAGAGCGCGGGAACCGGGCGGGGAGGAGGAGACACAGG 1480
|||||
QY 301 CTTTGACCGCATAGTAACCTCTCGCTCGGTGCGAGCGCGGAATCTATAAGAGAACTAGTCCC 360
|||||
Db 1481 CTTTGACCGCATAGTAACCTCTCGCTCGGTGCGAGCGCGGAATCTATAAGAGAACTAGTCCC 1540
|||||
QY 361 GGCACAAACCCCGTAATTGCGAGCGAGAGTGGTGGGGCGGAGCCCGGAGCGGAGCC 420
|||||
Db 1541 GGCACAAACCCCGTAATTGCGAGCGAGAGTGGTGGGGCGGAGCCCGGAGCGGAGCC 1600
|||||
QY 421 GACCCCTTCTCTCCCGGCTGCGGCGAGGCGAGGGCGGGGAGCTC 463
|||||
Db 1601 GACCCCTTCTCTCCCGGCTGCGGCGAGGCGAGGGCGGGGAGCTC 1643
|||||

RESULT 4

AAE92831
ID AAE92831 standard; DNA; 183999 BP.
XX AC
XX AAE92831;
DI 17-MAY-2001 (first entry)
XX
DE Human ABC1 genomic DNA.
XX
KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX
OS Homo sapiens.
XX
PN WO200115676-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-IB01492.
XX
PR 01-SEP-1999; 99US-0151977.
PR 15-MAR-2000; 2000US-0526193.
PR 23-JUN-2000; 2000US-0213958.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON GENETICS INC.
XX
PI Bayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX
DI WPI; 2001-244356/25.
XX
DE Treating a lower than normal high density lipoprotein-cholesterol
PI (HDL-C) level, a higher than normal triglyceride level, or a
PT cardiovascular disease, by administering a compound that modulates LXR-
PT or RXR-mediated transcriptional activity -
XX
XX Claim 8; Fig 1; 317pp; English.
PS
PS The present invention relates to a method for treating a patient
CC diagnosed as having a lower than normal high density
CC lipoprotein-cholesterol (HDL-C) level, a higher than normal

triglyceride level, or a cardiovascular disease, involving administering a compound that modulates LXR- or RXR-mediated transcriptional activity or ABC1 expression or activity. The LXR gene product may be used in an assay to identify compounds useful for the treatment of a disease or condition lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

XX	Key	Location/Qualifiers			
PT	protein_bind	12..23		protein_bind	/bound_moiety= "DeltaEFl"
PT		/*tag= a			/*tag= z
PT	protein_bind	/bound_moiety= "LMO2COM/MYCD"		protein_bind	/bound_moiety= "PPAR"
PT		97..107			/*tag= aa
PT	protein_bind	/*tag= b		protein_bind	/bound_moiety= "PPAR"
PT		/bound_moiety= "DeltaEFl"			1685..1698
PT		110..125			/*tag= ab
PT	protein_bind	/*tag= c		protein_bind	/bound_moiety= "USF/NMYC/MYCMAX"
PT		/bound_moiety= "S8/NKX2.5"			1787..1797
PT	protein_bind	196..211		protein_bind	/*tag= ac
PT		/*tag= d			/bound_moiety= "DeltaEFl"
PT	protein_bind	/bound_moiety= "S8"		protein_bind	1809..1819
PT		228..237			/*tag= ad
PT	protein_bind	/*tag= e		protein_bind	/bound_moiety= "DeltaEFl"
PT		/bound_moiety= "GATA"			1822..1833
PT	protein_bind	399..410		protein_bind	/*tag= ae
PT		/*tag= f			/bound_moiety= "SRY"
PT	protein_bind	/bound_moiety= "IK2"		protein_bind	1840..1850
PT		412..420			/*tag= af
PT	protein_bind	/*tag= g		protein_bind	/bound_moiety= "AP1"
PT		/bound_moiety= "LYF1"			1942..1956
PT	protein_bind	528..539		protein_bind	/*tag= ag
PT		/*tag= h			/bound_moiety= "HNF3beta"
PT	protein_bind	/bound_moiety= "LMO2COM/MYOD/DeltaEFl"		protein_bind	1978..1985
PT		549..556			/*tag= ah
PT	protein_bind	/*tag= i		protein_bind	/bound_moiety= "NKX2.5"
PT		/bound_moiety= "LYF1"			2008..2016
PT	protein_bind	558..568		protein_bind	/*tag= ai
PT		/*tag= j			/bound_moiety= "PPAR/NKX2.5/PPAR"
PT	protein_bind	/bound_moiety= "DeltaEFl"		protein_bind	2019..2024
PT		590..596			/*tag= aj
PT	protein_bind	/*tag= k		protein_bind	/bound_moiety= "PPAR/NKX2.5/PPAR"
PT		/bound_moiety= "LXX2.5"			2051..2059
PT	protein_bind	608..620		protein_bind	/*tag= ak
PT		/*tag= l			/bound_moiety= "GATA"
PT	protein_bind	/bound_moiety= "NFY/CAAT"		protein_bind	2104..2111
PT		708..715			/*tag= al
PT	protein_bind	/*tag= m		protein_bind	/bound_moiety= "SOX5"
PT		/bound_moiety= "MZFl"			2114..2152
PT	protein_bind	723..730		protein_bind	/*tag= am
PT		/*tag= n			/bound_moiety= "SRY/HFH/HNF3beta"
PT	protein_bind	/bound_moiety= "MZFl"		protein_bind	2221..2228
PT		771..785			/*tag= an
PT	protein_bind	/*tag= o		protein_bind	/bound_moiety= "MZFl"
PT		/bound_moiety= "HFH2/SRY/EV11"			2234..2249
PT	protein_bind	803..812		protein_bind	/*tag= ao
PT		/*tag= p			/bound_moiety= "IK2/NFKappaB/CREL"
PT	protein_bind	/bound_moiety= "CREBFl/VSp"		protein_bind	2259..2272
PT		831..837			/*tag= ap
PT	protein_bind	/*tag= r		protein_bind	/bound_moiety= "LMO2COM/GATA"
PT		/bound_moiety= "NKX2.5"			2289..2306
PT	protein_bind	1076..1089		protein_bind	/*tag= aq
PT		/*tag= s			/bound_moiety= "MZFl/SRY"
PT	protein_bind	/bound_moiety= "GATA"		protein_bind	2313..2318
PT		1173..1188			/*tag= ar
PT	protein_bind	/*tag= t		protein_bind	/bound_moiety= "PPAR"
PT		/bound_moiety= "LXRalpha/DeltaEFl"			2321..2326
PT	protein_bind	1189..1209		protein_bind	/*tag= as
PT		/*tag= u			/bound_moiety= "PPAR"
PT	protein_bind	/bound_moiety= "DeltaEFl/LYF1/IK2"		protein_bind	2335..2342
PT		1483..1491			/*tag= at
PT	protein_bind	/*tag= v		protein_bind	/bound_moiety= "MZFl"
PT		/bound_moiety= "AP4"			2361..2384
PT	protein_bind	1498..1514		protein_bind	/*tag= au
PT		/*tag= w			/bound_moiety= "HNF3beta/SRY/EV11"
PT	protein_bind	/bound_moiety= "LMO2-COM/MYOD/DeltaEFl/S47"		protein_bind	2426..2433
PT		1524..1545			/*tag= av
PT	protein_bind	/*tag= x		protein_bind	/bound_moiety= "MZFl"
PT		/bound_moiety= "ZID/DeltaEFl"			2442..2451
PT	protein_bind	1597..1607		protein_bind	/*tag= aw
PT		/*tag= y			/bound_moiety= "AP4"

```
protein_bind 2455..2466
FT          /*tag= ax
FT          /bound_moiety= "SRY"
protein_bind 2491..2498
FT          /*tag= ay
FT          /bound_moiety= "STAT"
protein_bind 2524..2534
FT          /*tag= az
FT          /bound_moiety= "STAT/PPAR"
protein_bind 2536..2541
FT          /*tag= ba
FT          /bound_moiety= "PPAR"
protein_bind 2589..2600
FT          /*tag= bb
FT          /bound_moiety= "AP2"
protein_bind 2510..2517
FT          /*tag= bc
FT          /bound_moiety= "MZF1"
protein_bind 2634..2648
FT          /*tag= bd
FT          /bound_moiety= "LMO2/CK/XYOD/E47"
protein_bind 2657..2672
FT          /*tag= be
FT          /bound_moiety= "RREB1"
protein_bind 2680..2698
FT          /*tag= bf
FT          /bound_moiety= "MZF1/CMYB"
protein_bind 2728..2740
FT          /*tag= bg
FT          /bound_moiety= "SP1/GC"
protein_bind 2743..2757
FT          /*tag= bh
FT          /bound_moiety= "USF/NMYC/ARNT"
protein_bind 2758..2773
FT          /*tag= bi
FT          /bound_moiety= "NFE2A1"
protein_bind 2774..2787
FT          /*tag= bj
FT          /bound_moiety= "XFD1/HFH"
protein_bind 2794..2806
FT          /*tag= bk
FT          /bound_moiety= "GC/SP1/MZF1"

Query Match      80.4%; Score 372.4; DB 24; Length 2510;
Best Local Similarity 98.5%; Pred. No. 7.7e-85;
Matches 388; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 COTTGGCTGCCGGGAACGTGGACTAGAGAGTCTGCGGCGCAGCCCGAGCCCGAGCGCTTC 60
DB 2517 COTTGGCTGCCGGGAACGTGGACTAGAGAGTCTGCGGCGCAGCCCGAGCGCTTC 2576

QY 61 CGGCGCTTTAGCGCGGGCGCGCGCGGGGAAGGGGAGCGAGCGCGGACCCCTAA 120
DB 2577 CGGCGCTTTAGCGCGGGCGCGCGCGGGGAAGGGGAGCGAGCGCGGACCCCTAA 2636

QY 121 GACACTGCTGTACCTCCACCCCAACCCACCC-----ACCTCCCCCACTCCCTAGA 175
DB 2637 GACACTGCTGTACCTCCACCCCAACCCCAACCCCACTCCCTCCCACTCCCTAGA 2696

QY 176 TGTGTGTTGGGGGGCTGAACGTGCGCCGTTTAAAGGGGGGGGGCGCGGTCCACGTGCTTT 235
DB 2697 TGTGTGTTGGGGGGCTGAACGTGCGCCGTTTAAAGGGGGGGGGCGCGGTCCACGTGCTTT 2756

QY 236 CTGCTGTGCTGAAGTACATTAACAGAGCGCGGGAAGGGGGGGGAGGAGGAGGAGC 295
DB 2757 CTGCTGTGCTGAAGTACATTAACAGAGCGCGGGAAGGGGGGGGAGGAGGAGGAGC 2816

QY 296 ACAGGCTTTGACGATAGTAAGCTCTGCGCTCGGTGACCGCAATCTATAAAGGAACATA 355
DB 2817 ACAGGCTTTGACGATAGTAAGCTCTGCGCTCGGTGACCGCAATCTATAAAGGAACATA 2876

QY 356 GTCCCCGGCAAAACCCCGTAATTGGAGCGAGAG 389
DB 389 GTCCCCGGCAAAACCCCGTAATTGGAGCGAGAG
```

```
Db 2877 GTCCGCGCAAAACCCCGTAATTGGAGCGAGAG 2910

RESULT 7
AAC09615
ID AAC09615 standard; cDNA; 227 BP.
XX AAC09615;
XX AAC09615;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 13690.
DE Humar; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID 13690; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from
mRNAs encoding secreted proteins. No ORF has yet been conclusively
identified within the present sequence. The 5' ESTs were prepared from
total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
sequences usually correspond mainly to the 3' untranslated region (UTR)
of the mRNA because they are often obtained from oligo-dT primed cDNA
libraries. Such ESTs are not well suited for isolating cDNA sequences
derived from the 5' ends of mRNAs and even in those cases where longer
cDNA sequences have been obtained, the full 5' UTR is rarely included.
5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
in diagnostic, forensic, gene therapy and chromosome mapping procedures.
They are used to obtain upstream regulatory sequences and to design
expression and secretion vectors.

Query Match      25.8%; Score 119.4; DB 21; Length 227;
Best Local Similarity 99.2%; Pred. No. 5.3e-21;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 343 ATAAAGGAAGTACTGTCGGCGCAAAACCCCGTAATTGGAGCGAGTCACTGGGCGGG 402
DB 1 ATAAAGGAAGTACTGTCGGCGCAAAACCCCGTAATTGGAGCGAGTCACTGGGCGGG 60

QY 403 GACCCGCGAGCGGAGCGACCCCTCTCTCCCGGCTCTCCGCGAGCGGCGGCGGAGCT 462
DB 61 GACCCGCGAGCGGAGCGACCCCTCTCTCCCGGCTCTCCGCGAGCGGCGGCGGAGCT 120

QY 463 C 463
DB 121 C 121

RESULT 8
```


CC erythematodes) is claimed. Modulation of ABC1 transcripts or
CC proteins by antisense or ribozyme technology or RNA decoys is also
CC claimed.
XX
CC
SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;
Query Match: 21.2%; Score 98; DP 22; Length 7260;
Best Local Similarity 100.0%; Pred. No. 2.5e-15;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 366 AAACCCCGTAATTCGAGCGAGAGTGTAGTGGCGCGGACCCGAGCGAGCGACCC 425
DB 1 AAACCCCGTAATTCGAGCGAGAGTGTAGTGGCGCGGACCCGAGCGAGCGACCC 60
QY 426 TTCCTCTCCCGGCTCGCGCAGGCGAGGCGCGGGAGCTC 463
DB 51 TTCCTCTCCCGGCTCGCGCAGGCGAGGCGCGGGAGCTC 92
RESULT 10
AAD37268
ID AAD37268 standard; DNA; 221 BP.
XX
AC AAD37268;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human ABC1 gene exon 1A.
XX
KW Human; ATP-binding cassette 1; ABC1 gene regulation; atherosclerosis;
KW cholesterol metabolism; hypercholesterolaemia; antisense therapy;
KW exon 1A; ds.
XX
XX Homo sapiens.
OS
XX WO200183746-A2.
PN
XX 08-NOV-2001.
PD
XX 02-MAY-2001; 2001WO-EP05488.
PF
XX 02-MAY-2000; 2000US-201280P.
PR
XX (AVET) AVENTIS PHARMA SA.
PA
XX Rosier-Montus M, Prades C, Lemoine C, Naudin L, Deneffe P;
PI Brewer B, Duveger N, Remaley A, Sanzamarina-Pojc S;
XX
DR WPI; 2002-154404/20.
XX
PT Isolated nucleic acid useful for modifying the ATP-binding cassette 1
PT (ABC1) and screening for candidate modulatory compounds or substances
PT
XX
PS Claim 4; Page 132; 152pp; English.
XX
CC The invention relates a nucleic acid which is capable of regulating the
CC transcription of human ATP-binding cassette 1 (ABC1) gene, which is a
CC causal gene for pathologies linked to a dysfunctioning of cholesterol
CC metabolism, including diseases such as atherosclerosis. Polynucleotides
CC of the invention are used to screen candidate molecules or substances
CC that are capable of modulating the transcription of the ABC1 gene. They
CC are used in antisense therapy. Compositions comprising sequences of the
CC invention are used to treat hypercholesterolaemia and atherosclerosis.
CC The present sequence is human ABC1 gene exon 1A.
XX
SQ Sequence 221 BP; 44 A; 62 C; 73 G; 42 T; 0 other;
Query Match 19.7%; Score 91; DP 24; Length 221;
Best Local Similarity 100.0%; Pred. No. 8.2e-14;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 373 GTAATTCGAGCGAGAGTGTAGTGGCGCGGACCCGAGCGAGCGACCCCTTCTC 432

DB 1 GTAATTCGAGCGAGAGTGTAGTGGCGCGGACCCGAGCGAGCGACCCCTTCTC 60
QY 433 CGGCGCTCGCGCAGGCGAGGCGCGGGAGCTC 463
DB 61 CGGCGCTCGCGCAGGCGAGGCGCGGGAGCTC 91
RESULT 11
AAH07432
ID AAH07432 standard; cDNA; 736 BP.
XX
AC AAH07432;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:4267.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183757.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES -NST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 4267; 2537pp - CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 3'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 3'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH24446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 19.3%; Score 89.4; DB 22; Length 736;
Best Local Similarity 98.9%; Pred. No. 2.6e-13; Indels 0; Gaps 0;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 373 GTAATTGCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGAGCCGACCTTCTCTC 432
DB 1 GTAGTTGCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGAGCCGACCTTCTCTC 60

QY 433 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 463
DB 61 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 91

RESULT 12
AAH18606
ID AAH18606 standard; cDNA; 1556 BP.
XX AC AAH18606;
XX AC AAH18606;
XX DT 26-JUN-2001 (first entry)
XX DE Human cDNA sequence SEQ ID NO:18808.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PE 28-JUL-2000; 2000EP-0116126.
XX PR 29-JUL-1999; 99JP-0248036.
XX PR 27-AUG-1999; 99JP-0300253.
XX PR 11-JAN-2000; 2000JP-0118776.
XX PR 02-MAY-2000; 2000JP-0183767.
XX PR 09-JUN-2000; 2000JP-0241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;
Query Match 19.3%; Score 89.4; DB 22; Length 1556;
Best Local Similarity 98.9%; Pred. No. 2.9e-13;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 373 GTAATTGCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGAGCCGACCTTCTCTC 432
DB 1 GTAGTTGCGAGCGAGTGTAGTGGCGCGGACCCGAGAGCGAGCCGACCTTCTCTC 60

QY 433 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 463
DB 61 CCGGGCTGCGGCGAGGCGAGGCGGGGAGCTC 91

RESULT 13
AAH04729
ID AAH04729 standard; cDNA; 763 BP.
XX AC AAH04729;
XX DT 26-JUN-2001 (first entry)
XX DE Human cDNA clone (5'-primer) SEQ ID NO:1564.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PE 28-JUL-2000; 2000EP-0116126.
XX PR 29-JUL-1999; 99JP-0248036.
XX PR 27-AUG-1999; 99JP-0300253.
XX PR 11-JAN-2000; 2000JP-0118776.
XX PR 02-MAY-2000; 2000JP-0183767.
XX PR 09-JUN-2000; 2000JP-0241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
PS Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 763 BP; 137 A; 205 C; 250 G; 158 T; 3 other;

Query Match 19.2%; Score 89; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 AATTCGAGCGAGAGTGTGGCGCGGACCGGAGCGGAGCGGAGCTC 434

Db 1 AATTCGAGCGAGAGTGTGGCGCGGACCGGAGCGGAGCGGAGCTC 60

QY 435 GGGCTCGGCGAGCGGAGCGGAGCGGAGCTC 463

Db 61 GGGCTCGGCGAGCGGAGCGGAGCGGAGCTC 89

RESULT 14

AAH17451

ID AAH17451 standard; cDNA; 1750 BP.

XX AC AAH-7451;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:16905.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074517-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama I, Wakamatsu A, Nagai K, Otsuki T;

XX XX WPI; 2001-318749/34.

XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs.

XX PS Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;

Query Match 19.2%; Score 89; DB 22; Length 1750;

Best Local Similarity 100.0%; Pred. No. 3.3e-13;

Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 AATTCGAGCGAGAGTGTGGCGCGGACCGGAGCGGAGCGGAGCTC 434

Db 1 AATTCGAGCGAGAGTGTGGCGCGGACCGGAGCGGAGCGGAGCTC 60

QY 435 GGGCTCGGCGAGCGGAGCGGAGCGGAGCTC 463

Db 61 GGGCTCGGCGAGCGGAGCGGAGCGGAGCTC 89

RESULT 15

AAK51683

ID AAK51683 standard; cDNA; 7281 BP.

XX AC AAK51683;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 223.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04092.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Zhang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX WPI; 2001-476283/51.

XX DR P-PSDB; AAK78550.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy.

XX PS Claim 1; Page 1086-1096; 6221pp; English.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 16:38:05 ; Search time 1007.79 Seconds
(without alignments)
10155.006 Million cell updates/sec
Title: US-09-596-141c-3_COPY_1292_1643
Perfect score: 352
Sequence: 1 ggaacctaagacaccctctg.....cagggcagggcggggcgctc 352

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054540 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pl.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	352	100.0	1167	9	HS252201 Homo sapi
2	352	100.0	1643	6	AX060715 Sequence
3	352	100.0	1643	6	AX060894 Sequence
4	352	100.0	96717	9	AL359182 Human DNA
5	352	100.0	175064	2	AC012230 Homo sapi
6	348.8	99.1	1167	9	AF238623sl
7	339.4	96.4	69570	2	AC021246 Homo sapi
8	336	95.5	183999	6	AX092589 Sequence
9	335.4	95.3	3231	6	AX351029 Sequence
10	335.4	95.3	149034	9	AF275948 Homo sapi
11	333.8	94.8	201144	9	AF287262 Homo sapi
12	271.6	77.2	69570	2	AC021246 Homo sapi
13	244.4	69.4	2893	6	AX351031 Sequence
14	227	64.5	90698	2	AC021345 Homo sapi
15	173	49.1	697	9	AF258627 Homo sapi
16	155.4	44.1	186889	2	AL807243 Mus muscu
17	155.4	44.1	278572	10	AF287263
18	133.2	43.5	145833	2	AC125637
19	98	27.8	7260	6	AX253452 Sequence
20	91	25.9	221	6	AX351032 Sequence
21	89.4	25.4	1556	9	AK024328 Homo sapi
22	89	25.3	1750	9	AK022254 Homo sapi
23	75	21.3	9854	6	AX127831 Sequence
24	75	21.3	9854	6	AX139818 Sequence
25	67	19.0	10442	6	AX060713 Sequence
26	67	19.0	10442	6	AX060892 Sequence
27	67	19.0	10442	9	AF285167 Homo sapi
28	60	17.0	10474	6	AX060719 Sequence
29	60	17.0	10474	6	AX060721 Sequence
30	60	17.0	10474	6	AX060898 Sequence
31	60	17.0	10474	6	AX060900 Sequence
32	43.6	12.4	170792	2	AC112455 Rattus no
33	43.2	12.3	188069	2	AC099251 Rattus no
34	41.4	11.8	136799	2	AC110334 Rattus no
35	40.6	11.5	3867	1	AY033407 Myxococcu
36	40.4	11.5	183022	2	AC109103 Rattus no
37	40.2	11.4	182996	9	AC027682 Homo sapi
38	39.6	11.2	195020	9	AF429315 Homo sapi
39	39.2	11.1	162811	2	AC095897 Rattus no
40	39	11.1	207959	2	AC129157 Rattus no
41	38.8	11.0	138151	2	AC094070 Rattus no
42	38.8	11.0	150994	2	AC111413 Rattus no
43	38.6	11.0	146606	2	AC111683 Rattus no
44	38.4	10.9	143330	2	AC094396 Rattus no
45	38.4	10.9	180731	2	AC103182 Rattus no

ALIGNMENTS

RESULT 1
HSA252201
LOCUS Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, 5'UTR and promoter region.
DEFINITION HSA252201 1167 bp DNA linear PRI 10-APR-2001
ACCESSION AJ252201
VERSION AJ252201.1 GI:12053757
KEYWORDS ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Porsch-Ozcurumez,M., Langmann,T., Heimerl,S., Borsukova,H., Kaminski,W.E., Drobnik,W., Honer,C., Schumacher,C. and Schmitz,G.

TITLE The zinc finger protein 202 (ZNF202) is a transcriptional repressor of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene expression and a modulator of cellular lipid efflux

JOURNAL J. Biol. Chem. 276 (15), 12427-12433 (2001)

MEDLINE 21192304

PUBMED 11279031

REFERENCE 2 (bases 1 to 1167)

AUTHORS Porsch-Oezueruemez M.K., Institute for Direct Submission

TITLE Submitted (03-JAN-2000)

JOURNAL Clinical Chemistry, University of Regensburg, Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY

FEATURES

source

1..1167

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="9q22-31"

/cell_type="leukocyte"

gene 1..1167

/gene="ABC-1"

promoter 1..942

/gene="ABC-1"

/function="cholesterol efflux regulatory protein"

TATA_signal 896..900

/gene="ABC-1"

intron 1148..1167

/gene="ABC-1"

/function="cholesterol efflux regulatory protein"

BASE COUNT 278 a 315 c 327 g 247 t

ORIGIN

Query Match 100.0%; Score 352; DB 9; Length 1167;

Best Local Similarity 100.0%; Pred. No. 2.5e-78;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACCTTAGACACCTGCTGTACCTCCACCCACCCACCCACCCACCTCCCTCCCACTCC 60

|||||

Db 666 GGACCTTAGACACCTGCTGTACCTCCACCCACCCACCCACCCACCTCCCTCCCACTCC 725

QY 61 TAGATGTGCTGGGGCTGAAGCTGCGCCGTTTAAAGGGGGGGGGGGGGGGGGGGTCCACGTG 120

|||||

Db 726 TAGATGTGCTGGGGCTGAAGCTGCGCCGTTTAAAGGGGGGGGGGGGGGGGGGGTCCACGTG 785

QY 121 CTTTCTGTAGTGAACACATATAACAGAGCGCGGGAACGGGGGGGGGGAGGAGGA 180

|||||

Db 786 CTTTCTGTAGTGAACACATATAACAGAGCGCGGGAACGGGGGGGGGGAGGAGGA 845

QY 181 GAGCAGAGCTTTGACCGATAGTAACTCTGCGCTGCGTGCAGCGCAATCTATAAAGGA 240

|||||

Db 846 GAGCAGAGCTTTGACCGATAGTAACTCTGCGCTGCGTGCAGCGCAATCTATAAAGGA 505

QY 241 ACTAGTCCCGGCAAAACCCCTAATTGGAGAGAGTGAAGTGGGGGGGGGGGGGGGGGGAG 300

|||||

Db 906 ACTAGTCCCGGCAAAACCCCTAATTGGAGAGAGTGAAGTGGGGGGGGGGGGGGGGGGAG 965

QY 301 AGCCGAGCGACCCCTCTCTCCCGGCTCGGCGAGGCGAGGGGGGGGAGCTC 352

|||||

Db 966 AGCCGAGCGACCCCTCTCTCCCGGCTCGGCGAGGCGAGGGGGGGGAGCTC 1017

RESULT 2

AX060715

LOCUS 1643 bp DNA linear PAT 22-JAN-2001

DEFINITION Sequence 3 from Patent WO0078972.

ACCESSION AX060715

VERSION AX060715.1 GI:1240604

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1643)

AUTHORS Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.

TITLE ATP binding cassette transporter protein abci polypeptides

JOURNAL Patent: WO 0078971-A 3 28-DEC-2000;

CV THERAPEUTICS, INC. (US)

FEATURES

source

1..1643

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t

ORIGIN

Query Match 100.0%; Score 352; DB 6; Length 1643;

Best Local Similarity 100.0%; Pred. No. 2.4e-78;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACCTTAGACACCTGTGTACCTCCACCCACCCACCCACCCACCTCCCTCCCACTCC 60

|||||

Db 1292 GGACCTTAGACACCTGTGTACCTCCACCCACCCACCCACCCACCTCCCTCCCACTCC 1351

QY 61 TAGATGTGCTGGGGCTGAAGCTGCGCCGTTTAAAGGGGGGGGGGGGGGGGGGGTCCACGTG 120

|||||

Db 1352 TAGATGTGCTGGGGCTGAAGCTGCGCCGTTTAAAGGGGGGGGGGGGGGGGGGGTCCACGTG 1411

QY 121 CTTTCTGTAGTGAACACATATAACAGAGCGCGGGAACGGGGGGGGGGAGGAGGA 180

|||||

Db 1412 CTTTCTGTAGTGAACACATATAACAGAGCGCGGGAACGGGGGGGGGGAGGAGGA 1471

QY 181 GAGCAGAGCTTTGACCGATAGTAACTCTGCGCTGCGTGCAGCGCAATCTATAAAGGA 240

|||||

Db 1472 GAGCAGAGCTTTGACCGATAGTAACTCTGCGCTGCGTGCAGCGCAATCTATAAAGGA 1531

QY 241 ACTAGTCCCGGCAAAACCCCTAATTGGAGAGAGTGAAGTGGGGGGGGGGGGGGGGGGAG 300

|||||

Db 1532 ACTAGTCCCGGCAAAACCCCTAATTGGAGAGAGTGAAGTGGGGGGGGGGGGGGGGGGAG 1591

QY 301 AGCCGAGCGACCCCTCTCTCCCGGCTCGGCGAGGCGAGGGGGGGGAGCTC 352

|||||

Db 1592 AGCCGAGCGACCCCTCTCTCCCGGCTCGGCGAGGCGAGGGGGGGGAGCTC 1643

Db 1352 TAGATGCTGCTGGGGGCTGAACGCTCCCGCTTAAAGGGGGGGGGCCCGGCTCACGTG 1411
 Qy 121 CTTCCTGCTGAGTGAATACATAACAGAGCCCGGACCGGGGGGAGAGGGA 180
 |||||
 Db 1412 CTTCCTGCTGAGTGAATACATAACAGAGCCCGGACCGGGGGGAGAGGGA 1471
 Qy 181 GAGCAGAGCTTGAACCATAGTAACTCTCGCTGCTGAGCGGGAATCTATAAAGGA 240
 |||||
 Db 1472 GAGCAGAGCTTGAACCATAGTAACTCTCGCTGCTGAGCGGGAATCTATAAAGGA 1531
 Qy 241 ACTAGTCCCGGCAAAACCCCGTAAITCGGAGCGAGAGTGAAGTGGGGCGGACCCGAG 300
 Db 1532 ACTAGTCCCGGCAAAACCCCGTAAITCGGAGCGAGAGTGAAGTGGGGCGGACCCGAG 1591
 Qy 301 AGCGAGCGGACCCCTTCTCTCCCGGGCTGCGGACAGGCGAGGGGGGAGCTC 352
 |||||
 Db 1592 AGCGAGCGGACCCCTTCTCTCCCGGGCTGCGGACAGGCGAGGGGGGAGCTC 1543

RESULT 4

AL359182/c AL359182 96717 bp DNA linear PRI 11-JAN-2002
 LOCUS Human DNA sequence from clone RP11-217B7 on chromosome 9, comp.ete
 DEFINITION sequence.
 ACCESSION AL359182
 VERSION AL359182.20 GI:18151453
 KEYWORDS HTG.
 SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 96717)

REFERENCE

Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jan 15, 2002 this sequence version replaced gi:18121468.

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em, EMBL; Sw, SW-SSRPRO; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
 database can be found at

<http://www.sanger.ac.uk/projects/Celegans/wormpep> This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
 RP11-217B7 is from the library RP11-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-217B7. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-217B7 is at 96717 in this
 sequence. The true left end of clone RP11-222F10 is at 72980 in
 this sequence. The true right end of clone RP11-31J20 is at 2000 in
 this sequence.

Location/Qualifiers

1..96717

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-217B7"
 /clone_id="RP11-11.1"
 3238..3278
 /note="Sequence from AF275948 sequenced by National
 Institutes of Health, National Heart, Lung and Blood
 Institute, Bethesda, MD 20892, USA."
 84249..84273
 /note="Sequence from overlapping clone RP11-122F10
 (AC026643). Assembly confirmed by restriction digest."
 92050..92163
 /note="Sequence from reads from a short insert library
 derived from a single pUC clone. Restriction digest data
 confirm the assembly."
 92411..92557
 /note="Sequence from reads from a short insert library
 derived from a single pUC clone. Restriction digest data
 confirm the assembly."
 27673 a 21138 c 20380 g 27526 t
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 352; DB 9; Length 96717;
 Best Local Similarity 100.0%; Pred. No. 1.6e-78;
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGACCCTAAGACACCTGCTGACCTCCACCCACCCACCCACCTCCCTCCCACTCC 60
 |||||
 Db 17518 GGACCCTAAGACACCTGCTGACCTCCACCCACCCACCCACCTCCCTCCCACTCC 17459
 Qy 61 TAGATGCTGCTGGGGGCTGAACGCTCCCGCTTAAAGGGGGGGGGCTCCACCTG 120
 |||||
 Db 17458 TAGATGCTGCTGGGGGCTGAACGCTCCCGCTTAAAGGGGGGGGGCTCCACCTG 17399
 Qy 121 CTTTCTGCTGAGTGAATACATAACAGAGCCCGGACCGGGGGGAGAGGGA 180
 |||||
 Db 17398 CTTTCTGCTGAGTGAATACATAACAGAGCCCGGACCGGGGGGAGAGGGA 17339
 Qy 181 GAGCAGAGCTTGAACCATAGTAACTCTCGCTGCTGAGCGGGAATCTATAAAGGA 240
 |||||
 Db 17338 GAGCAGAGCTTGAACCATAGTAACTCTCGCTGCTGAGCGGGAATCTATAAAGGA 17279
 Qy 241 ACTAGTCCCGGCAAAACCCCGTAAITCGGAGCGAGAGTGAAGTGGGGCGGACCCGAG 300
 |||||
 Db 17278 ACTAGTCCCGGCAAAACCCCGTAAITCGGAGCGAGAGTGAAGTGGGGCGGACCCGAG 17219
 Qy 301 AGCGAGCGGACCCCTTCTCTCCCGGGCTGCGGACAGGCGAGGGGGGAGCTC 352
 |||||
 Db 17218 AGCGAGCGGACCCCTTCTCTCCCGGGCTGCGGACAGGCGAGGGGGGAGCTC 17167

RESULT 5

AC012230

LOCUS

DEFINITION

Accession

Version

Keywords

Source

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)

1 (bases 1 to 175064)


```

misc_feature 17301..20131
              /note="assembly_fragment"
misc_feature 20232..22587
              /note="assembly_fragment"
misc_feature 22688..25707
              /note="assembly_fragment"
misc_feature 25808..28184
              /note="assembly_fragment"
misc_feature 28285..31338
              /note="assembly_fragment"
misc_feature 31439..34299
              /note="assembly_fragment"
misc_feature 34200..38318
              /note="assembly_fragment"
misc_feature 38419..42835
              /note="assembly_fragment"
misc_feature 42936..45448
              /note="assembly_fragment"
misc_feature 45549..48116
              /note="assembly_fragment"
misc_feature 48217..52618
              /note="assembly_fragment"
misc_feature 52719..56592
              /note="assembly_fragment"
misc_feature 56893..59635
              /note="assembly_fragment"
misc_feature 59736..63661
              /note="assembly_fragment"
misc_feature 63762..68437
              /note="assembly_fragment"
misc_feature 68538..71458
              /note="assembly_fragment
              close.end:sp6
              vector_side:left"
misc_feature 71559..76888

Query Match      100.0%; Score 352; DB 2; Length 175064;
Best Local Similarity 100.0%; Pred. No. 1.5e-78;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACCTTAAGACACCTGCTGTACCTCCACCCGCCACCCACCCACCTCCGCCCACTCCC 60
Db 3401 GGACCTTAAGACACCTGCTGTACCTCCACCCGCCACCCACCCACCTCCGCCCACTCCC 3460

Qy 61 TAGATGCTGCTGGCGGCTGAACCTGCGCGGTTTAAAGGGGGGGGGGGGGGGGGGGGG 120
Db 3461 TAGATGCTGCTGGCGGCTGAACCTGCGCGGTTTAAAGGGGGGGGGGGGGGGGGGGGG 3520

Qy 121 CTTTCTGCTGAGTGAACCTACATAACAGAGCGCGGGGAGGAGGAGGGA 180
Db 3521 CTTTCTGCTGAGTGAACCTACATAACAGAGCGCGGGGAGGAGGAGGGA 3580

Qy 181 GAGCACAGGCTTTGACCGATAGTAACTTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAG 240
Db 3581 GAGCACAGGCTTTGACCGATAGTAACTTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAG 3640

Qy 241 ACTAGTCCCGGCAAAACCCCGTATTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAGCG 300
Db 3641 ACTAGTCCCGGCAAAACCCCGTATTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAGCG 3700

Qy 301 AGCCGAGCGGACCCCTCTCTCCCGGGCTCGCGCAGGCGAGGCGGAGGAGCTC 352
Db 3701 AGCCGAGCGGACCCCTCTCTCCCGGGCTCGCGCAGGCGAGGCGGAGGAGCTC 3752

RESULT 6
LOCUS AF258623S1 1167 bp DNA linear PRI 23-JUN-2000
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,
            promoter and exon 1.
ACCESSION AF258623
VERSION AF258623.2 GI:8677405
KEYWORDS

```

```

SEGMENT 1 of 4
SOURCE Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
          Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of ABC1 gene 5' end: additional peptide sequence,
          promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 224 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
          Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,
          University of California, San Francisco, 505 Parnassus Avenue, San
          Francisco, CA 94143-0130, USA
REFERENCE 3 (bases 1 to 1167)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
          Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,
          University of California, San Francisco, 505 Parnassus Avenue, San
          Francisco, CA 94143-0130, USA
REMARK Sequence update by submitter
COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.
FEATURES
          Location/Qualifiers
            source
              1..1167
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="9"
                /map="9q31"
                /gene="ABCA1"
                /gene="ABCA1"
                /number=1
            promoter
              224..844
            exon
              845..1147
                /gene="ABCA1"
                /number=1
BASE COUNT 278 a 313 c 328 g 248 t
ORIGIN
          Query Match      99.1%; Score 348.8; DB 9; Length 1167;
          Best Local Similarity 99.4%; Pred. No. 1.6e-77;
          Matches 350; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGACCTTAAGACACCTGCTGTACCTCCACCCGCCACCCACCTCCGCCCACTCCC 60
Db 566 GGACCTTAAGACACCTGCTGTACCTCCACCCGCCACCCACCTCCGCCCACTCCC 725

Qy 61 TAGATGCTGCTGGCGGCTGAACCTGCGCGGTTTAAAGGGGGGGGGGGGGGGGGGGGG 120
Db 726 TAGATGCTGCTGGCGGCTGAACCTGCGCGGTTTAAAGGGGGGGGGGGGGGGGGGGGG 785

Qy 121 CTTTCTGCTGAGTGAACCTACATAACAGAGCGCGGGGAGGAGGAGGGA 180
Db 786 CTTTCTGCTGAGTGAACCTACATAACAGAGCGCGGGGAGGAGGAGGGA 845

Qy 181 GAGCACAGGCTTTGACCGATAGTAACTTCGCGAGCGAGAGTGAAGTGGGGCGGAGCG 240
Db 846 GAGCACAGGCTTTGACCGATAGTAACTTCGCGAGCGAGAGTGAAGTGGGGCGGAGCG 905

Qy 241 ACTAGTCCCGGCAAAACCCCGTATTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAG 300
Db 906 ACTAGTCCCGGCAAAACCCCGTATTCGCGAGCGAGAGTGAAGTGGGGCGGAGCGGAG 965

Qy 301 AGCCGAGCGGACCCCTCTCTCCCGGGCTCGCGCAGGCGAGGCGGAGGAGCTC 352
Db 966 AGCCGAGCGGACCCCTCTCTCCCGGGCTCGCGCAGGCGAGGCGGAGGAGCTC 1017

RESULT 7
LOCUS AC021246 69570 bp DNA linear HTG 13-JUL-2000

```



```

DEFINITION Homo sapiens clone RP11-IN10, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC021246
VERSION AC021246.2 GI:9119882
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69570)
Barren,B., Linton,L., Nusbaum,C., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,I., Bookhvalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Collangelo,K., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,A., Jones,C., Kan,J., Karatas,A., Klein,J.,
Landers,T., Leoczky,J., Levine,R., Liu,C., Liu,G., Locke,R.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Melgrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705871.
All repeats were identified using RepeatMasker:
Sait, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W13R
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2512
Center clone name: 1_N_10
-----
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 871: contig of 871 bp in length
872 971: gap of 100 bp
972 1834: contig of 863 bp in length
1835 1935: gap of 100 bp
1935 2804: contig of 870 bp in length
2805 2904: gap of 100 bp
2905 3745: contig of 841 bp in length
3745 3845: gap of 100 bp
3845 4696: contig of 851 bp in length
4697 4796: gap of 100 bp
4797 5640: contig of 844 bp in length
5641 5740: gap of 100 bp
5741 6540: contig of 800 bp in length
6541 6640: gap of 100 bp
6641 7509: contig of 869 bp in length
7510 7609: gap of 100 bp
*
* 8710 8473: contig of 970 bp in length
* 8480 8579: gap of 100 bp
* 8580 9430: contig of 851 bp in length
* 9431 9530: gap of 100 bp
* 9531 10376: contig of 846 bp in length
* 10377 10476: gap of 100 bp
* 10477 11322: contig of 846 bp in length
* 11323 11422: gap of 100 bp
* 11423 12302: contig of 830 bp in length
* 12303 12402: gap of 100 bp
* 12403 13280: contig of 878 bp in length
* 13281 13380: gap of 100 bp
* 13381 14241: contig of 861 bp in length
* 14242 14341: gap of 100 bp
* 14342 15196: contig of 855 bp in length
* 15197 15296: gap of 100 bp
* 15297 16123: contig of 827 bp in length
* 16124 16223: gap of 100 bp
* 16224 17072: contig of 849 bp in length
* 17073 17172: gap of 100 bp
* 17173 18041: contig of 869 bp in length
* 18042 18141: gap of 100 bp
* 18142 19009: contig of 868 bp in length
* 19010 19109: gap of 100 bp
* 19110 19966: contig of 857 bp in length
* 19967 20066: gap of 100 bp
* 20067 20921: contig of 855 bp in length
* 20922 21021: gap of 100 bp
* 21022 21866: contig of 844 bp in length
* 21866 21965: gap of 100 bp
* 21966 22832: contig of 867 bp in length
* 22833 22932: gap of 100 bp
* 22933 23780: contig of 848 bp in length
* 23781 23880: gap of 100 bp
* 23881 24733: contig of 853 bp in length
* 24734 24833: gap of 100 bp
* 24834 25670: contig of 837 bp in length
* 25671 25770: gap of 100 bp
* 25771 26621: contig of 851 bp in length
* 26622 26721: gap of 100 bp
* 26722 27576: contig of 855 bp in length
* 27577 27676: gap of 100 bp
* 27677 28532: contig of 856 bp in length
* 28533 28632: gap of 100 bp
* 28633 29492: contig of 860 bp in length
* 29493 29592: gap of 100 bp
* 29593 30455: contig of 863 bp in length
* 30456 30555: gap of 100 bp
* 30556 31410: contig of 855 bp in length
* 31411 31510: gap of 100 bp
* 31511 32368: contig of 858 bp in length
* 32369 32468: gap of 100 bp
* 32469 33312: contig of 844 bp in length
* 33313 33412: gap of 100 bp
* 33413 34268: contig of 856 bp in length
* 34269 34368: gap of 100 bp
* 34369 35204: contig of 836 bp in length
* 35205 35304: gap of 100 bp
* 35305 36156: contig of 852 bp in length
* 36157 36256: gap of 100 bp
* 36257 37128: contig of 872 bp in length
* 37129 37228: gap of 100 bp
* 37229 38083: contig of 855 bp in length
* 38084 38183: gap of 100 bp
* 38184 39031: contig of 848 bp in length
* 39032 39131: gap of 100 bp
* 39132 40006: contig of 875 bp in length
* 40007 40106: gap of 100 bp
* 40107 40967: contig of 861 bp in length
* 40968 41067: gap of 100 bp
* 41068 41913: contig of 846 bp in length
* 41914 42013: gap of 100 bp
* 42014 42824: contig of 811 bp in length

```

```
* 42825 42924: gap of 100 bp
* 42925 43776: contig of 852 bp in length
* 43777 43876: gap of 100 bp
* 43877 44752: contig of 876 bp in length
* 44753 44852: gap of 100 bp
* 44853 44724: contig of 872 bp in length
* 45725 45824: gap of 100 bp
* 45825 45643: contig of 819 bp in length
* 45644 46743: gap of 100 bp
* 46744 47599: contig of 856 bp in length
* 47600 47698: gap of 100 bp
* 47700 48551: contig of 852 bp in length
* 48552 48651: gap of 100 bp
* 48652 49485: contig of 834 bp in length
* 49486 49585: gap of 100 bp
* 49586 50440: contig of 855 bp in length
* 50441 50540: gap of 100 bp
* 50541 51404: contig of 864 bp in length
* 51405 51504: gap of 100 bp
* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55228: contig of 861 bp in length
* 55230 55329: gap of 100 bp
* 55330 55197: contig of 888 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match          96.4%; Score 339.4; DB 2; Length 69570;
Best Local Similarity 99.4%; Pred. No. 2.4e-75;
Matches 351; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GGACCTTAAGACACCTGCTGTACCTCCACCCACCCACCCACCTCCCGCCCAACT-CC 59
|||||
Db 41303 GGACCTTAAGACACCTGCTGTACCTCCACCCACCCACCCACCTCCCGCCCAACTGCC 41362
|||||

Qy 60 CTAGATGTCTCGTGGGGGCTGAAGCTGGCGGCTTTAAGGGCGGGCGGCGGCTCCAGT 119
|||||
Db 41363 CTATATGTCTCGTGGCGGCTGAAGCTGGCGGCTTTAAGGGCGGGCGGCTCCAGT 41422
|||||

Qy 120 GCTTCTCTCTGAGTGACTGAACATACATAACAGAGCGCGGAGCGGGCGGGAGGAGG 179
|||||
Db 41423 GCTTCTCTCTGAGTGACTGAACATACATAACAGAGCGCGGAGCGGGCGGGAGGAGG 41482
|||||

Qy 180 AGAGCAGAGGCTTTGACCGCAAGTAACTCTCGGCTCGGTGACCGCCGAATCTATAAAGG 239
|||||
Db 41483 AGAGCAGAGGCTTTGACCGCAAGTAACTCTCGGCTCGGTGACCGCGGAATCTATAAAGG 41542
|||||
```

```
Qy 240 AACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGGGCGGAGCCGCA 299
|||||
Db 41543 AACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGGGCGGAGCCGCA 41602
|||||

Qy 300 GAGCGAGCGGACCTTCTCICCGGGCTCGGCGAGGCGAGGCGGGGAGCTC 352
|||||
Db 41603 GAGCGAGCGGACCTTCTCICCGGGCTCGGCGAGGCGAGGCGGGGAGCTC 41655
|||||

RESULT 8
AX092589 183999 bp DNA linear PAT 21-MAR-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0115676.
ACCESSION AX092589
VERSION AX092589.1 GI:13444647
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 183999)
AUTHORS Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M.
TITLE Compositions and methods for modulating hdl cholesterol and
triglyceride levels
JOURNAL Patent: WO 0115676-A 1 08-MAR-2001;
FEATURES
Location/Qualifiers
source
1..183999
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 49549 a 37944 c 41170 g 54950 t 386 others
ORIGIN
Query Match 95.5%; Score 336; DB 6; Length 183999;
Best Local Similarity 98.3%; Pred. No. 1.6e-74;
Matches 351; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 1 GGACCTTAAGACACCTGCTGTACCTCCA-----CCCCACCCACCCACCTCCCGCCCAA 55
|||||
Db 28524 GGACCTTAAGACACCTGCTGTACCTCCANNNNCCACCACCCACCTCCCGCCCAA 28583
|||||

Qy 56 CTCCTAGATGTGCTGGGGGCTGAAGCTGGCGGCTTTAAGGGCGGGCGGCGGCTCC 115
|||||
Db 28584 CTCCTAGATGTGCTGGGGGCTGAAGCTGGCGGCTTTAAGGGCGGGCGGCGGCTCC 28643
|||||

Qy 116 ACGTGCCTTCTGCTGAGTGAACACTACATAACAGAGGCGGGAACGGGCGGGGAGG 175
|||||
Db 28644 ACGTGCCTTCTGCTGAGTGAACACTACATAACAGAGGCGGGAACGGGCGGGGAGG 28703
|||||

Qy 176 AGGAGACACAGGCTTTCACCGGATAGTAACCTCTGCGCTCGGTGACCGCAATCTATAA 235
|||||
Db 28704 AGGAGACACAGGCTTTCACCGGATAGTAACCTCTGCGCTCGGTGACCGCAATCTATAA 28763
|||||

Qy 236 AAGAACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGGGCGGAGCC 295
|||||
Db 28754 AAGAACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGGGCGGAGCC 28823
|||||

Qy 296 CGCAGAGCGAGCGGACCCCTTCTCTCCGGGCTGCGGAGCGGCGAGGCGGGGAGCTC 352
|||||
Db 28824 CGCAGAGCGAGCGGACCCCTTCTCTCCGGGCTGCGGAGCGGCGAGGCGGGGAGCTC 28880
|||||

RESULT 9
AX351029 3231 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 1 from Patent WO0183746.
ACCESSION AX351029
VERSION AX351029.1 GI:18616385
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE 1
AUTHORS Rosier-Montus M.F., Prades C., Lemoine C., Naudin L., Deneffe P.,
Brewer B., Duverger N., Remaley A. and Santamarina-Fojo S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 1 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES Location/Qualifiers
Source 1. 3231
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 809 a 773 c 876 g 773 t
ORIGIN
Query Match 95.3%; Score 335.4; DB 6; Length 3231;
Best Local Similarity 98.3%; Pred No. 3.4e-74;
Matches 351; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
QY 1 GGACCTTAGACACCGTGTACCTCCAC-----CCCCACCCACCCACCTCCCCCCAA 55
|||||
Db 2628 GGACCTTAGACACCGTGTACCTCCACCCACCCACCCACCCACCTCCCCCCAA 2687
QY 56 CTCCTAGATGTCGTCGGCGGCTGAACGTGCGCCGCTTAAGGCGCGCGCCGCTCC 115
|||||
Db 2688 CTCCTAGATGTCGTCGGCGGCTGAACGTGCGCCGCTTAAGGCGCGCGCCGCTCC 2747
QY 116 ACCTGCTTCTCTGCTAGTGAAGTGAACATAACAGAGCGCGGAGCGGGGAGG 175
|||||
Db 2748 ACCTGCTTCTCTGCTAGTGAAGTGAACATAACAGAGCGCGGAGCGGGGAGG 2807
QY 176 AGGAGACACAGGCTTACCGATAGTAACTCTCGGTCGGTCGAGCGCGAATCTATAA 235
|||||
Db 2808 AGGAGACACAGGCTTACCGATAGTAACTCTCGGTCGGTCGAGCGCGAATCTATAA 2867
QY 236 AAGGACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGCGCGGACC 295
|||||
Db 2868 AAGGACTAGTCCCGCAAAACCCCGTAATTCGAGCGAGAGTGAGTGGCGCGGACC 2927
QY 296 CGCAGAGCGAGCGAGCGCTTCTCCCGGCTCGGCGAGCGAGCGCGGAGGCTC 352
|||||
Db 2928 CGCAGAGCGAGCGAGCGCTTCTCCCGGCTCGGCGAGCGAGCGCGGAGGCTC 2984
RESULT 10
LOCUS AF275948 149034 bp DNA linear PRI 17-JUL-2000
DEFINITION Homo sapiens ABCA1 (ABCA1) gene, complete cds.
ACCESSION AF275948
VERSION AF275948.1 GI:9247085
SOURCE Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 149034)
AUTHORS Santamarina-Fojo S., Peterson K., Knapper C., Qiu Y., Freeman L.,
Cheng J.F., Osorio J., Remaley A., Yang X.P., Haudenschield C.,
Prades C., Chimini G., Blackmon E., Francois T., Duverger N.,
Rubin E.M., Rosier M., Deneffe P., Fredrickson D.S. and Brewer H.B.
Jr.
TITLE Complete genomic sequence of the human ABCA1 gene: analysis of the
human and mouse ATP-binding cassette A promoter
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (14), 7987-7992 (2000)
MEDLINE 20345099
PUBMED 10864428
REFERENCE 2 (bases 1 to 149034)
AUTHORS Santamarina-Fojo S., Peterson K.M., Knapper C.L., Freeman J.A.,
Remaley A.T., Yang X.-P., Haudenschield C.C., Blackmon E.B.,
Francois T.L. and Brewer H.B. Jr.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2000) Molecular Disease Branch, National
Institutes of Health, National Heart, Lung and Blood Institute,
Bethesda, MD 20892, USA

Location/Qualifiers
1. 149034
/organism="Homo sapiens"
/db_xref="taxon:9606"
1454. 148034
/gene="ABCA1"
Join(1454. 1674,25831. 25989,40385. 40478,45012. 45153,
46423. 46541,67729. 67850,70831. 71007,83960. 84052,
89010. 89250,91962. 92101,92433. 92549,96758. 96955,
97702. 97907,98428. 98604,100391. 100613,102360. 102581,
103642. 103846,104951. 105064,106862. 107033,
108023. 108154,109460. 109602,109806. 109943,
110646. 110866,112124. 112196,113183. 113385,
115053. 115101,115287. 115420,116806. 116954,
118804. 118728,119984. 120062,123094. 123283,
124804. 124898,126208. 126240,127363. 127468,
128943. 129017,129539. 129708,130939. 131116,
133114. 133229,133341. 133485,134526. 134649,
135736. 135865,136130. 136250,137038. 137100,
138008. 138114,140468. 140610,140982. 141116,
142060. 142163,142646. 142738,143397. 143640,
144581. 148034)
/gene="ABCA1"
/product="ABCA1"
Join(1454. 1674,25831. 25923)
2406. 2678
/rpt_family="Alu"
4240. 4509
/rpt_family="Alu"
4998. 5270
/rpt_family="Alu"
5563. 5839
/rpt_family="Alu"
6615. 6877
/rpt_family="Alu"
8800. 9046
/rpt_family="Alu"
9998. 10279
/rpt_family="Alu"
11859. 12133
/rpt_family="Alu"
12810. 12902
/rpt_family="Alu"
15220. 15399
/rpt_family="Alu"
15508. 15787
/rpt_family="Alu"
18600. 18860
/rpt_family="Alu"
20950. 21206
/rpt_family="Alu"
25039. 25314
/rpt_family="Alu"
Join(25924. 25989,40385. 40478,45012. 45153,46423. 46541,
67729. 67850,70831. 71007,83960. 84052,89010. 89250,
91962. 92101,92433. 92549,96758. 96955,97702. 97907,
98428. 98604,100391. 100613,102360. 102581,103642. 103846,
104951. 105064,106862. 107033,108023. 108154,
109460. 109602,109806. 109943,110646. 110866,
112124. 112196,113183. 113385,115053. 115101,
115297. 115410,116806. 116954,118604. 118728,
119964. 120062,123094. 123283,124804. 124898,
126208. 126240,127363. 127468,128943. 129017,
129539. 129708,130939. 131116,133114. 133229,
133485. 133485,134526. 134649,135736. 135865,
136130. 136250,137038. 137100,138008. 138114,
140469. 140610,140982. 141116,142060. 142163,
142646. 142738,143397. 143640,144581. 144721)
/gene="ABCA1"
/codon_start=1
/product="ABCA1"
/protein_id="AAF66276.1"

```

/db_xref="CI:3247086"
/translation="MACPQURLLLWKXNLTRRRQTCQLLEVAWPFILILTSYV
SYYPQEHCFHPNKAHPFSACTLPWQOICINANNPQFRYPRTQGEAPGVGNFNKSTIV
ARLSPKSRVLLYSOKDTSMKDMKVLRIQOIKKSSGNKLQDFLVDVNDFFSGVLVH
NLSLPKSPDQKMLRADVTLHKVPLQGVOLHLLSLCNGSSKDEMIQLQDQSVSLQCLIP
REKLAARERVLRSNNWDLKPLILRTINSTSPPPSKELARATKILLHSGLTQAQLSFSMR
SWSNRQEVMEFLVNVNNSSTQIIVQSVRISVCGHPGGGLKLSLNYEDNNIKALF
GGNGTEDEAETFYDNSTIYCNLMKNLSEPLSRIWKALKPLLYGLIKYLTIDTPAT
COLGDAWNTQELAEHFDHDEKMEELSPKTIWFENMSQNDQVRLMDDSRDNDHWEVL
POVMDLTAQDIQVAFIAKHPEDDVQSNGSVYTWREAFNETQAIRISREMECVNLN
KLEPATIEMVLLIKSMELLDNRKRWAGIVFGIIPGSELPHVHYKIRIDYNKINERT
NKIKQVDPGPRADPDMRYVWVGGRAYLDQVQOALIVLGTPEKTKYVQWQQPY
PCYDDIDILRYNMSRSMPLFWLAMIYSAVNIKGIYKIEKARLKEIMRIMGLNLSILW
PSWFISLIPLLVSAGLLVILKLGNIPLSDPSVVFVLSVFAVVIILOCFILSTLF
SRANLACAGGIIIFTIIPVYLCAVMDQYVGTFLKIFAXLSPVAFSGCEYFALFPE
YGGVQVNDNFESPVEDGFNLTIYSMMKLDFTLXGVMTWIEAVFPQOYGIIPRPW
YFPCTKSYWFGDESKHSGNSKRISETCMEDEPETHLKLGVSIQNLVAVYRQGMKV
AVDGLANFYEGQVTSLEGHNGACKTMTSILGTLPFTPTGSTAVILGKDTRSMSTIR
QNLVGPPOHNVLFOMLVEEHIHVFYARLKGISEKHVKAEMEOHALDVLGIPSSKLKKT
SOLSGMOKLSVALAFVGGSKVILDEPTAGUDVPYSRRGIWELLKAYIROGRIITLST
HMBEADVLGDRIALIISHGKLUCCVSGSSLFLKNQLOGTGYLTLVKKDVESSLSCRNSS
STVSLKKKDESVOSSDAGLGSDESDTITDVSAINLIRKHVSAREVEDIGHEL
TYVPLAEAKGAFVPELHEIDRDLSDGLISSYGETTLEEIPLKVAESGVDATIS
DGTLPARNRKKQKQSCLRPFPTEDDAADPNDSIDIPESRETDLGSGMDGKGSYOVK
GKWTIQQOQFVALLNKRLLIARSKSGPFAQIVLPAVFCIALYVFLVLPFGKYPSLE
LQPMYNDQVILFVNDAPEDGTIELNALTKDPFGCTRCMEGNPTIEDTPCQAGEEZW
TTAPVOTIMDLPONGWNTQNSPACQCSDDIKKMLVCPGPGAGLPPQRRQNTA
DILODLTGRNLSDYLVYTVQIILAKLNKIWNEFRYGGFSLGVSNTQALPFSQVYN
DAXKOMKHLKALASDADRFLNLSKSLFKNQLOGTGYLTLVKKDVESSLSCRNSS
NAILRANLQKGENPSHYGITAIFNHPNLTKQOISEVAXMTSYDVLVSGICVIFAMGFPV
PASVFWPLTQBRVSKAKHLQFISGKPEVIYWLNSFWMDMGNVYVPAVLVIIITCQO
KSYVSSNTLPYALLLILLYGWSIPTLMPASFEVKIPSTAVVYLSVNLITIGINGSVA
IFVLELDTLKNLNNIILKSVFLIFPHCLGRLIDMKVNOQAMADALFEGENRVPVS
PLSDVLGRNLFANAGVGVFELITVLOYRFFIPRPPVANAKLSPUNDEDEDVRRERQ
RILDQGGONILEIKELTKIYRRKXFAVRDRCVGPFGCFLGVLVNGAKGSTNRM
LTCGTTVTRGDAPVNXNLSUNHEVHONGYCPQDAITELLITGREHVEFFALLRGV
PEKRVGVEWAIKRLGLVYGEKAGYAGNSGNNRKLSTAMALITGPPVYFLDEPTTG
MDKAPARLEALC-SVYKGRSVYLTSHMSECEALCIRMAIWNVRFRCLGSLGVQHLK
NRFQDGTIVYRIAGSNPDLPKPVODTFGLAFPGSVXKEKHNNMLQYOLPSSLSLARI
PSILSQSKKRLHIEDYSVSQTTLDQEVFNFAKQSDGDDHLKDLSLHKNQIVVDVAVLT
SFLQDEKVESYV"
29916..30180
/rpt_family="Alu"
/rpt_family="Alu"
31050..31314
/rpt_family="Alu"
31692..31973
/rpt_family="Alu"
35167..36440
/rpt_family="Alu"
37617..37892
/rpt_family="Alu"
37999..38166
/rpt_family="Alu"
38172..38269
/rpt_family="Alu"
38995..39271
/rpt_family="Alu"
41547..41565
/gene="ABCA1"
/note="A009X28"
41593..41835
/rpt_family="Alu"
41657..41678
/gene="ABCA1"
/note="A009X28"
42320..42584
/rpt_family="Alu"
44349..44626
/rpt_family="Alu"
48473..48669
/rpt_family="Alu"
50547..50614
/rpt_family="Alu"
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
STS
repeat_region
STS
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 Source
 gene
 mRNA
 exon
 exon
 CDS

21251004
 11352567
 2 (bases 1 to 201144)
 Qiu,Y., Cavelier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
 Direct Submission
 Submitted (13-JUL-2000) Genome Science Department, Lawrence
 Berkeley National Laboratory, 1 Cyclotron Rd., MS 84-171, Berkeley,
 CA 94720, USA

Location/Qualifiers
 1..201144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 33331..181457
 /gene="ABCA1"
 join(33931..34151,58319..58476,72893..72986,77520..77661,
 78931..79049,101048..101169,104152..104328,117287..117379,
 123239..123269,125285..125424,125754..125870,
 130079..130276,131023..131228,131749..131925,
 133708..133930,135677..135898,136959..137163,
 138268..138381,140179..140350,141340..141471,
 142777..142919,143123..143260,143963..144183,
 145441..145513,146657..146859,148527..148775,
 148775..148884,150280..150426,152078..152202,
 153438..153536,156568..156757,158278..158372,
 159682..159714,160837..160942,162417..162491,
 163013..163182,164413..164590,166588..166703,
 166815..166959,168505..168628,169715..169844,
 170109..170229,171017..171079,171987..172093,
 174448..174589,174961..175095,176039..176142,
 176625..176717,177376..177619,178560..181457)
 /gene="ABCA1"
 /product="ATP-binding cassette 1 sub-family A member 1"
 33931..34151
 /gene="ABCA1"
 /number=1
 58319..58476
 /gene="ABCA1"
 /number=2
 join(58411..58476,72893..72986,77520..77661,78931..79049,
 101048..101169,104152..104328,117287..117379,
 123239..123269,125285..125424,125754..125870,
 130079..130276,131023..131228,131749..131925,
 133708..133930,135677..135898,136959..137163,
 138268..138381,140179..140350,141340..141471,
 142777..142919,143123..143260,143963..144183,
 145441..145513,146657..146859,148527..148775,
 148775..148884,150280..150426,152078..152202,
 153438..153536,156568..156757,158278..158372,
 159682..159714,160837..160942,162417..162491,
 163013..163182,164413..164590,166588..166703,
 166815..166959,168505..168628,169715..169844,
 170109..170229,171017..171079,171987..172093,
 174448..174589,174961..175095,176039..176142,
 176625..176717,177376..177619,178560..178700)
 /gene="ABCA1"
 /codon_start=1
 /product="ATP-binding cassette 1 sub-family A member 1"
 /protein_id="AAK43526.1"
 /db_xref="GI:13875613"
 /translation="MACWEQLRLIKLWKLIRRRQICQLLEAVNPLFIFILISVRL
 SYPEQHECPNPKAMPSATLPWQGIICNANPCPRPIPGAPGVGNFKSIY
 ARLSDFARLLISOKDTSMKMRVLRLLQIKASSNKLQDLVDNETFSGLYH
 NLSLPKTEVLMRADYILHVFQYQLHLSICNGSKSEMIQLQVSELCGLP
 KEKAAARVLRMRDLKPIRLTNSPFSKELAPATKILHSLGTLAQELFSMR
 SWSMDRQVFLINVSSSTOIVCAVSRVCGHPGGGKIKSLNEDNNKALF
 GNGTEAEFTYDNSTPYCNDLKNLLESSPLSRILWALKPLLVGLIYPTPAI
 RQVAENVKTOELAVFHLGDMWELSPKLTWFMENQENDLVRMLDSRDNDHFW
 QQLGLDQADIVAFKHPEDVSSNGSVYTWEAFNETNOAIRTSREWEVNLK
 KLEPTATVRLINKSMELLDERKFWAGIVFTGTPGSELPHVVKYKTRMDIDNVERT
 NKIKDQWDPGRADPDMRYVGGFAVLQDVQCAIIRVLTGTEKTKYVMOOMPY
 PCYDDIDFLVNSRMP-FMPLAWISYAVIIGIYKEKARLKEETMELGDSILW
 FSWRTSSILPLVNSGLLVIILKGNLLPYSDPSVVFVFAVVTILQCFLISTLF
 SRNLAAACGGIITFLPLPVLVAVWODYVGT-LKIPASLLSPVAFGEGCEQFALFE

ECGIGVQNDLFEPSVEEDGFNLTTTSVSMFLNFTLYGVMTWYIEAVFPQXQIRPW
 YFCTKSYWFESEDEKSHPGSNQKRISEICMEEPHLKLGAVISONLVKVRDGMKV
 AVGLALNFYEQGIITFGLNGAGKTITMSILTLGTPPTSGTAYILGKDIRSDMSIR
 ONLGYCPQHNVLFDMLTVEHIEWEVARLKLKKVKAEMKALDVGLPSSKLKSKI
 SOLSGMORLKVAFVGSXVWILDETPAGVDPYSRRGIWELLIKYRQGRITILLST
 HMDRADVLGDRIALISHGKLCVSSLEFNOLGTYLTVKKDSESLSSCRNS
 TTVLPYAAKGAFAVELFHEIDRLSDLSISYSETTLEEFILKVAEALVSDIQREL
 DTLPLARRNRAGFKQKSLRPTEDDAADNDSDIDPSRETDLLSGMDGKGSQVK
 GWKLTQQOQFVLLWKLRLTARSRKGFAGIYLPVAVFCIALVPSLIVPFGKYSLE
 LOPMYNEOYTVSNDAPEDTGTLELLNALTQDPGEGTRCMENPIPTPQACGEEW
 TTAPVPTIMDLFQNCNWTQNPSPACQSSDKIRKMLPCCPGAGGLPPQPKONTA
 DIQDITGRNISDYLVKTYITIAKSLNKKIWNFEFRYGGFSLGVSNTQALPPOEVN
 DAIKMKRLHLAKDSSADRLNSLGRMTGLDIRNNKVFNNKWEHAISSFLAVIN
 NAILRANLQKGVNSHYGITAFNHPILNTKQSLSEVALMTSDVLSVCSVIFAMSFV
 PASVYVFLQERVSRAKHQFISGVKPVYIWLNSFVDMCNVWVATLVIIFICFOQ
 KSVSTNLPVLALLLLYGSITPLMYPASVEKTPSPAYVLSVNLFCINGSVA
 TTVLELFTDNKLNINDILKSVFLIPFHECLGRGLLDKVKYNAMADALERGFENFVS
 PLSWDLVGNLFAMAVEGVFFLTVLQYRFRPRPVNAKLSPLNDEDEVRKRO
 RLDDGGQNDLLEIKELTKIYRRKRKPAVDRICVGPGECEGLLVKSGAGKSTFKM
 LTGDTVTGRDAFLKKNSTLSNIHEVHONMGYCPQDAITELLTGRHVEFTALLRGV
 PEKEVGKGVGAIRKILGLVKYKTAGNYSNGKRLKLTAMALIGGPPVYFDEPTIG
 MDPKARFLWNCALSVKSGRSVLTSHSMECEALCTPMATWNRGRRCGLSGYHLK
 NREGDGYTVIWSIAGSNPDLKPVQDFGLAFPGSVLKEKHNMLQYQLPSSUSSLARI
 PSILSQSKRLHIEDYSVSQTLQDQVFNFAKQDQDDDELKDLSLHKNQTVVDAVL
 SFLQDEKVKRESTV"
 72893..72986
 /gene="ABCA1"
 /number=3
 77520..77661
 /gene="ABCA1"
 /number=4
 78931..79049
 /gene="ABCA1"
 /number=5
 101048..101169
 /gene="ABCA1"
 /number=6
 104152..104328
 /gene="ABCA1"
 /number=7
 117287..117379
 /gene="ABCA1"
 /number=8
 123239..123269
 /gene="ABCA1"
 /number=9
 125285..125424
 /gene="ABCA1"
 /number=10
 125754..125870
 /gene="ABCA1"
 /number=11
 130079..130276
 /gene="ABCA1"
 /number=12
 131023..131228
 /gene="ABCA1"
 /number=13
 131749..131925
 /gene="ABCA1"
 /number=14
 133708..133930
 /gene="ABCA1"
 /number=15
 135677..135898
 /gene="ABCA1"
 /number=16
 136959..137163
 /gene="ABCA1"
 /number=17
 138268..138381
 /gene="ABCA1"

```

/exon=18
-40179..140350
/gene="ABCA1"
/exon=19
141340..141471
/gene="ABCA1"
/exon=20
142777..142919
/gene="ABCA1"
/exon=21
143123..143260
/gene="ABCA1"
/exon=22
143963..144183
/gene="ABCA1"
/exon=23
145441..145513
/gene="ABCA1"
/exon=24
146657..146859
/gene="ABCA1"
/exon=25
148527..148575
/gene="ABCA1"
/exon=26
148771..148884
/gene="ABCA1"
/exon=27
150280..150428
/gene="ABCA1"
/exon=28
152078..152202
/gene="ABCA1"
/exon=29
153438..153536
/gene="ABCA1"
/exon=30
156558..156757
/gene="ABCA1"

```

```

Query Match      94.8%; Score 333.8; DB 9; Length 201144;
Best Local Similarity 98.0%; Pred. No. 5.6e-74;
Matches 350; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```

```

Qy 1 GGACCTAAGACACCTGTGTACCTCCAC-----CCACACCCACCCACCTCCGCCCAA 55
Db 33665 GGACCTAAGACACCTGTGTACCTCCACCCACCCACCCACCCACCTCCGCCCAA 33724
Qy 56 CTCCTAGATGTGCTGGCGGCTGACGTGCGCCGTTTAAAGGGCGGGCGCCCGGCTCC 115
Db 33725 CTCCTAGATGTGCTGGCGGCTGACGTGCGCCGTTTAAAGGGCGGGCGCCCGCTCC 33784
Qy 116 ACGTGCTTCTGCTGAGTGACTGAACACTACATAACAGAGGCGCGGAACGGCGGGGAGG 175
Db 33785 ACGTGCTTCTGCTGAGTGACTGAACACTACATAACAGAGGCGCGGAACGGCGGGAGG 33844
Qy 176 AGGAGACACAGCCTTGACCGATAGTAACTCTGCTCGCTCGTGACCCGAACTATPAA 235
Db 33845 AGGAGACACAGCCTTGACCGATAGTAACTCTGCTCGCTCGTGACCCGAACTATPAA 33904
Qy 236 AAGAACTAGTCCCGCAAAAACCCGTAATTGGAGGAGAGTGAAGTGGCGCGGGGACC 295
Db 33905 AAGAACTAGTCTGGCAAAAACCCCGTAATTGGAGGAGAGTGAAGTGGCGCGGGGACC 33964
Qy 296 CGCAGAGCCGACCGACCTTCTCTCCGGGCTGCGGAGGCGGAGGGCGGGAGCTC 352
Db 33965 CGCAGAGCCGACCGACCTTCTCTCCGGGCTGCGGAGGCGGAGGGCGGGAGCTC 34021

```

```

RESULT 12
AC021246/c
LOCUS
DEFINITION Homo sapiens clone RP11-IN10, LOW-PASS SEQUENCE SAMPLING.
          AC021246          69570 bp      DNA      linear      13-JUL-2000

```

AC021246

AC021246.2 GI:9.13882
 FIG: HTGS_PHASE0.
 Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 69570)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome, clone RP11-IN10

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 69570)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
 Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,S., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearrellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Seaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karas,A., Klein,J.,
 Landers,T., Lechoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 McSheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojatovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Thirell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6705871.

COMMENT

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WB3R

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2512

Center clone name: L_N10

* NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

```

1      871: contig of 871 bp in length
*
*
972    971: gap of 100 bp
*
*
1835   1834: contig of 863 bp in length
*
*
1935   1934: gap of 100 bp
*
*
2805   2804: contig of 870 bp in length
*
*
2805   2904: gap of 100 bp
*
*
2905   3745: contig of 841 bp in length
*
*
3746   3845: gap of 100 bp
*
*
3846   4596: contig of 851 bp in length
*
*
4697   4796: gap of 100 bp
*
*
4797   5640: contig of 844 bp in length
*
*
5641   5740: gap of 100 bp
*
*
5741   6540: contig of 800 bp in length
*
*
6541   6640: gap of 100 bp
*
*
6641   7509: contig of 869 bp in length
*
*
7510   7609: gap of 100 bp
*
*
7610   8479: contig of 870 bp in length
*

```



```
QY 271 AGCGAGTGTAGTGGCGGGGACCCGAGCGAGCGACCGTCTCTCTCCGGGCTGC 330
|||||
Db 67402 ACGAGAGTGTAGTGGCGGGGACCCGAGCGAGCGACCGTCTCTCTCCGGTCTGC 67343
QY 331 GCGAGGCGAGGCGGGGAGCTC 352
|||||
Db 67342 GACAGGACAGGCGGGGAGCTC 67321

RESULT 13
AX351031 AX351031 2893 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 3 from Patent WO0183746.
DEFINITION
ACCESSION AX351031
VERSION AX351031.1 GI:18616387
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,
Brewer, B., Duvergier, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 3 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
Location/Qualifiers
source 1..2893
/organism="Homo sapiens"
/ds_xref="taxon:9606"
BASE COUNT 745 a 681 c 769 g 698 t
ORIGIN

Query Match 69.4%; Score 244.4; DB 6; Length 2893;
Best Local Similarity 97.7%; Pred. No. 2.8e-51;
Matches 260; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 GGACCTTAAGACACCTGTCTACCTCCAC-----CCCCACCCACCCACCTCCCCCCAA 55
|||||
Db 2628 GGACCTTAAGACACCTGTCTACCTCCACCCACCCACCCACCTCCCCCCAA 2687
QY 56 CTCCTAGATGTCTGTGGGGGCTGAAGTCGCCGCTTAAGGGGCGGGCCCGGCTCC 115
|||||
Db 2688 CTCCTAGATGTCTGTGGGGGCTGAAGTCGCCGCTTAAGGGGCGGGCCCGGCTCC 2747
QY 116 ACCTGTTTCTGCTGACTGAAGTACATATAACAGAGCGCGGAGCGGGGCGAGG 175
|||||
Db 2748 ACCTGTTTCTGCTGACTGAAGTACATATAACAGAGCGCGGAGCGGGGCGAGG 2807
QY 176 AGGAGACACAGGCTTTCACCGATAGTAACCTTCGCTCGGTGAGCGCGGATCTATAA 235
|||||
Db 2808 AGGAGACACAGGCTTTCACCGATAGTAACCTTCGCTCGGTGAGCGCGGATCTATAA 2867
QY 236 AAGGAAGTGTCCCGCAAAACCCC 261
|||||
Db 2868 AAGGAAGTGTCCCGCAAAACCCC 2893

RESULT 14
AC021345/c AC021345 90698 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90698)
REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.
AUTHORS Homo sapiens, clone RP11-24J9
TITLE Unpublished
JOURNAL
```

REFERENCE
AUTHORS

2 (bases 1 to 90698)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavsky, L., Bocklender, B., Brown, A., Burkett, G., Castle, A.,
Choquet, Y., Collangelo, M., Collins, S., Collymore, A., Cocke, P.,
DeArelano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kahn, L., Karatas, A., Klein, J.,
Landers, T., Lechoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C.E., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, E., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced 91:6705761.

COMMENT

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4483

Center clone name: 24_J-9

* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Rups of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 910: contig of 910 bp in length
911 1010: gap of 100 bp
1011 1873: contig of 863 bp in length
1874 1973: gap of 100 bp
1974 2824: contig of 851 bp in length
2825 2924: gap of 100 bp
2925 3902: contig of 878 bp in length
3903 3902: gap of 100 bp
3903 4816: contig of 914 bp in length
4817 4916: gap of 100 bp
4917 5759: contig of 843 bp in length
5760 5859: gap of 100 bp
5860 6764: contig of 905 bp in length
6765 6864: gap of 100 bp
6865 7747: contig of 883 bp in length
7748 7847: gap of 100 bp
7848 8755: contig of 908 bp in length
8756 8855: gap of 100 bp
8856 9753: contig of 898 bp in length
9754 9853: gap of 100 bp
9854 10757: contig of 904 bp in length
10758 10857: gap of 100 bp
10858 11732: contig of 875 bp in length
11733 11832: gap of 100 bp
11833 12739: contig of 907 bp in length
12740 12840: gap of 100 bp
12840 13710: contig of 871 bp in length
13711 13810: gap of 100 bp


```
* 13811 14684: contig of 874 bp in length
14685 14784: gap of 100 bp
14785 15662: contig of 878 bp in length
15663 15762: gap of 100 bp
15763 16677: contig of 915 bp in length
16678 16777: gap of 100 bp
16778 17678: contig of 901 bp in length
17679 17778: gap of 100 bp
17779 18675: contig of 901 bp in length
18680 18779: gap of 100 bp
18780 19632: contig of 853 bp in length
19633 19732: gap of 100 bp
19733 20634: contig of 902 bp in length
20635 20734: gap of 100 bp
20735 21620: contig of 886 bp in length
21621 21730: gap of 100 bp
21731 22573: contig of 859 bp in length
22580 22679: gap of 100 bp
22680 23568: contig of 889 bp in length
23569 23668: gap of 100 bp
23669 24554: contig of 886 bp in length
24555 24654: gap of 100 bp
24655 25521: contig of 867 bp in length
25522 25621: gap of 100 bp
25622 26487: contig of 866 bp in length
26488 26587: gap of 100 bp
26588 27464: contig of 877 bp in length
27465 27554: gap of 100 bp
27555 28466: contig of 902 bp in length
28467 28556: gap of 100 bp
28567 29464: contig of 898 bp in length
29465 29564: gap of 100 bp
29565 30447: contig of 883 bp in length
30448 30547: gap of 100 bp
30548 31453: contig of 906 bp in length
31454 31533: gap of 100 bp
31534 32452: contig of 899 bp in length
32453 32552: gap of 100 bp
32553 33447: contig of 895 bp in length
33448 33547: gap of 100 bp
33548 34435: contig of 888 bp in length
34436 34535: gap of 100 bp
34536 35433: contig of 898 bp in length
35434 35533: gap of 100 bp
35534 36440: contig of 907 bp in length
36441 36540: gap of 100 bp
36541 37422: contig of 882 bp in length
37423 37522: gap of 100 bp
37523 38402: contig of 880 bp in length
38403 38502: gap of 100 bp
38503 39360: contig of 878 bp in length
39361 39480: gap of 100 bp
39481 40407: contig of 927 bp in length
40408 40507: gap of 100 bp
40508 41405: contig of 898 bp in length
41406 41505: gap of 100 bp
41506 42359: contig of 894 bp in length
42400 42499: gap of 100 bp
42500 43357: contig of 858 bp in length
43358 43457: gap of 100 bp
43458 44356: contig of 899 bp in length
44357 44456: gap of 100 bp
44457 45325: contig of 869 bp in length
45326 45425: gap of 100 bp
45426 46305: contig of 880 bp in length
46306 46405: gap of 100 bp
46406 47302: contig of 897 bp in length
47303 47402: gap of 100 bp
47403 48293: contig of 891 bp in length
48294 48393: gap of 100 bp
48394 49280: contig of 887 bp in length
49281 49380: gap of 100 bp
49381 50250: contig of 870 bp in length
```

```
* 50251 50350: gap of 100 bp
50351 51209: contig of 859 bp in length
51210 51309: gap of 100 bp
51310 52221: contig of 912 bp in length
52222 52321: gap of 100 bp
52322 53205: contig of 884 bp in length
53206 53305: gap of 100 bp
53306 54223: contig of 920 bp in length
54224 54323: gap of 100 bp
54324 55206: contig of 881 bp in length
55207 55306: gap of 100 bp
55307 56165: contig of 859 bp in length
56166 56265: gap of 100 bp
56266 57124: contig of 859 bp in length
57125 57224: gap of 100 bp
57225 58121: contig of 897 bp in length
58122 58221: gap of 100 bp
58222 59075: contig of 854 bp in length
59076 59175: gap of 100 bp
59176 60088: contig of 883 bp in length
60089 60188: gap of 100 bp
60189 61067: contig of 909 bp in length
61068 61167: gap of 100 bp
61168 62051: contig of 884 bp in length
62052 62151: gap of 100 bp
62152 63022: contig of 871 bp in length
63023 63122: gap of 100 bp
63123 64026: contig of 904 bp in length
64027 64126: gap of 100 bp
64127 65015: contig of 889 bp in length
65016 65115: gap of 100 bp
65116 65975: contig of 860 bp in length
65976 66075: gap of 100 bp
66076 66954: contig of 879 bp in length
66955 67054: gap of 100 bp
67055 67947: contig of 893 bp in length
67948 68047: gap of 100 bp
68048 68934: contig of 887 bp in length
68935 69034: gap of 100 bp
69035 69910: contig of 875 bp in length
69911 70010: gap of 100 bp

Query Match 64.3% Score 227; Db 2; Length 90698;
Best Local Similarity 90.2%; Pred. No. 4.7e-47;
Matches 275; Conservative 0; Mismatches 27; Indels 3; Gaps 3;

QY 48 CCCCCCACTCCCTAGATGTGTCGTGGGGGCTGAACGTGCCCCGTTAAAGGCGCGGCC 107
Db 87745 CCCCACCTCCCACTCATAG-GTGTCTGGCGCTGAACGTGCGCGG-TTAAGGGCGCGGCC 87687

QY 108 CCGGCTCCAGTGGCTTCTGCTGAGTGACTCACTACATACACAGAGCGGGACCGGG 167
Db 87686 CCGGCTCCAGTGGCTTCTGCTGAGTGTGGAACTACATAACAGAG-CCCGGGAACCGGC 87628

QY 168 CCGGAGGAGGAGGAGACAGGCTTTCACGCAATAGTAACTCTGCTGCTGAGCGCA 227
Db 87627 CCGGAGGAGGAGGAGACAGGCTTTCACGCAATAGTAACTCTGCTGCTGAGCGCA 87568

QY 228 ATCTATAAAGGAAGTACTAGTCCCGCAAAACCCCGTAATTCGAGAGAGTGTAGTGGG 287
Db 87567 ATCTATAAAGGAAGTACTAGTCCCGCAAAACCCCGTAATTCGAGAGAGTGTAGTGGG 87509

QY 288 CCGGAGCCCGCAGAGCGGAGCGGACCGCTCTCTCCCCGGGTGGCGAGGAGCGCGGG 347
Db 87508 CCGGAGCCCGCAGAGCGGAGCGGACCGCTCTCTCCCCGGGTGGCGAGGAGCGCGGG 87449

QY 348 AGCTC 352
Db 87448 AGCTC 87444
```

RESULT 15
AF258627

LOCUS AF258627 697 bp mRNA linear PRI 11-MAY-2000
 DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
 partial cds.
 ACCESSION AF258627
 VERSION AF258627.1 GI:7769707
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 697)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J., and Kane,J.P.
 TITLE Analysis of ABCA1 gene 5' end: additional peptide sequence,
 promoter region, and four polymorphisms
 JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
 REFERENCE 2 (bases 1 to 697)
 AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
 Aouizerat,B.E., Fielding,C.J., and Kane,J.P.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
 University of California, San Francisco, 505 Parnassus Avenue, San
 Francisco, CA 94143-0130, USA
 FEATURES
 Location/Qualifiers
 1..697
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q31"
 /tissue_type="placenta"
 1..>697
 /gene="ABCA1"
 /note="ABCA1"
 396..>697
 /gene="ABCA1"
 /note="membrane-bound"
 /cgc_start=1
 /product="ATP binding cassette transporter 1"
 /protein_id="AAF69513.1"
 /db_xref="GI:7769708"
 /translation="MACWPOLRLIWKJFRRQTCOLLLEVAWPLFIFLILISVRL
 SYPPYEQHECHFPNKAMPSAGTIPWVQGLICNANNPCFRYPTEAPGVVGNFNKS"

BASE COUNT 152 a 138 c 190 g 156 t 1 others
 ORIGIN
 Query Match 49.1%; Score 173; DB 9; Length 697;
 Best Local Similarity 100.0%; Pred. No. 3.1e-33;
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 180 AGAGCACAGGCTTTGACCGATAGTAACTCTGCGCTGGTGCAGCCGCAATCTATAAAGG 239
 Db 1 AGAGCACAGGCTTTGACCGATAGTAACTCTGCGCTGGTGCAGCCGCAATCTATAAAGG 60
 QY 240 AACTAGTCCCGGCAAAACCCGTAATTGCGAGCGAGTGGGCGGCGGACCGCA 299
 Db 61 AACTAGTCCCGGCAAAACCCGTAATTGCGAGCGAGTGGGCGGCGGACCGCA 120
 QY 300 GAGCCGAGCGAGCCCTCTCTCCGGGCTGGGCGAGGCGGCGGAGCTC 352
 Db 121 GAGCCGAGCGAGCCCTCTCTCCGGGCTGGGCGAGGCGGCGGAGCTC 173

Search completed: February 4, 2003, 01:32:23
 Job time : 1336.79 secs